



GENOME ET VARIABILITE

La variabilité des génomes individuels humains

Partie 1 (2 x 2h, C. Borel, session 1 et 2 agenda)

Types et origine des variants

Partie 2 (2h, C. Borel, session 3 agenda)

Conséquences phénotypiques

Partie 3 (1h, C. Borel, session 4 agenda)

Source de la diversité humaine

HEREDITE

Les modes de transmission des maladies génétiques

Partie 1 (2h, M. Neerman-Arbez, session 5 agenda)

Partie 2 (1h, M. Neerman-Arbez, session 6 agenda)

Partie 3 (2h, C. Borel, session 7 agenda) Hérité complexe des maladies multifactorielles

Génétique et génomique du cancer (2h, T. Nospikel, session 8 agenda)

TD Génétique humaine (2h, M. Neerman-Arbez, C. Borel)

Iconographie

L'examen porte sur l'ensemble du cours !



EXEMPLE CLINIQUE



APPROCHE EXPERIMENTALE



QUESTION NON RESOLUE



POUR LES SPECIALISTES



BIEN CONNAITRE LES DETAILS

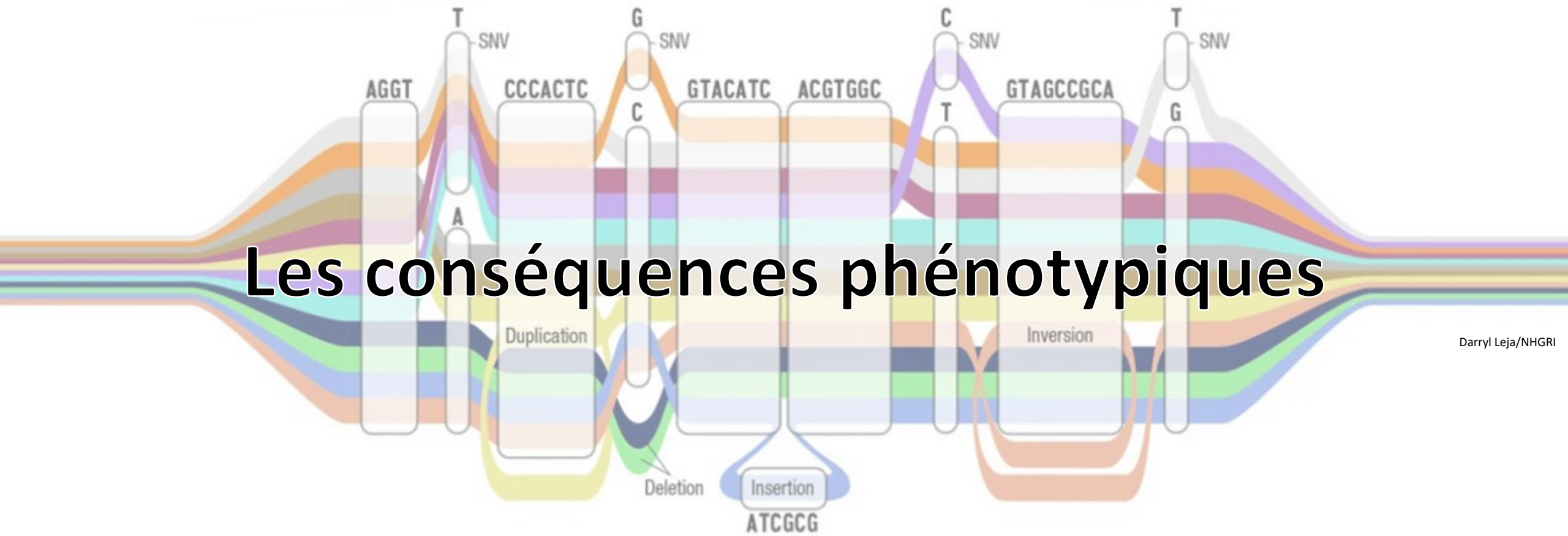
UNIQUEMENT PENDANT LE COURS



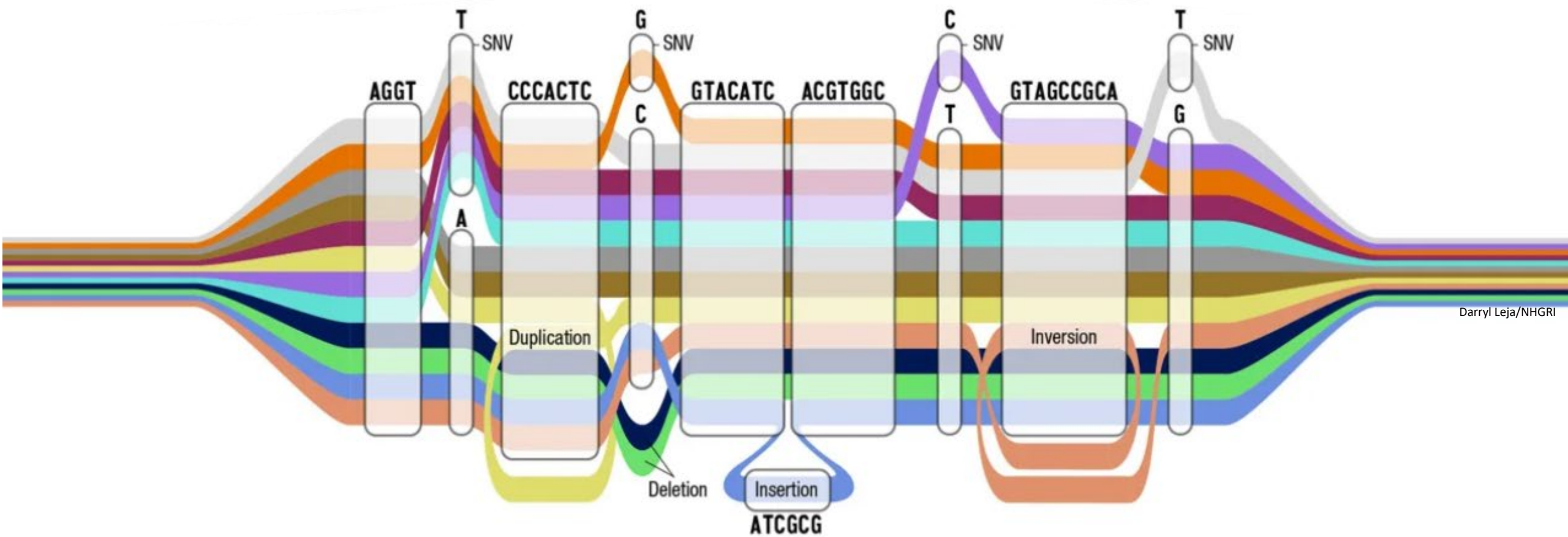
Site web:
<https://web.speakup.info/room/join/87493>
Numéro de salle speakup: 87493



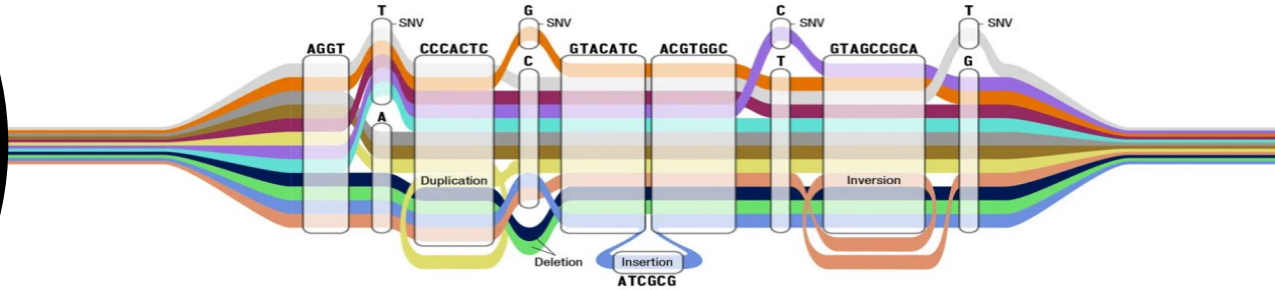
Les conséquences phénotypiques



Darryl Leja/NHGRI



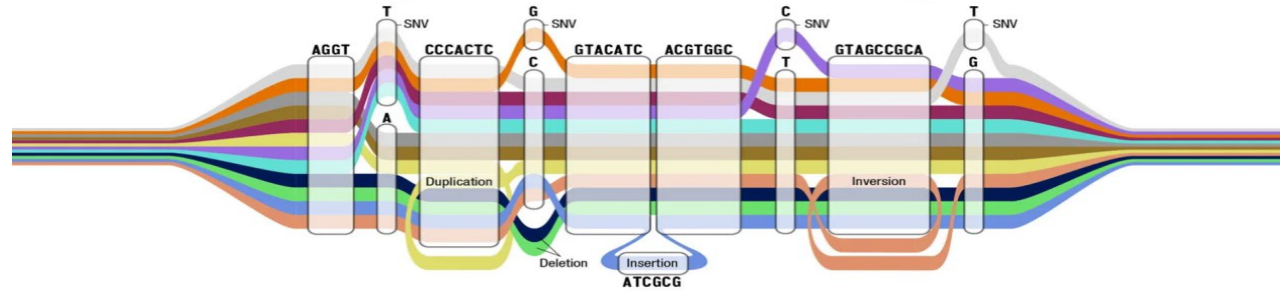
Darryl Leja/NHGRI



807 162 génomes
d'individus séquencés

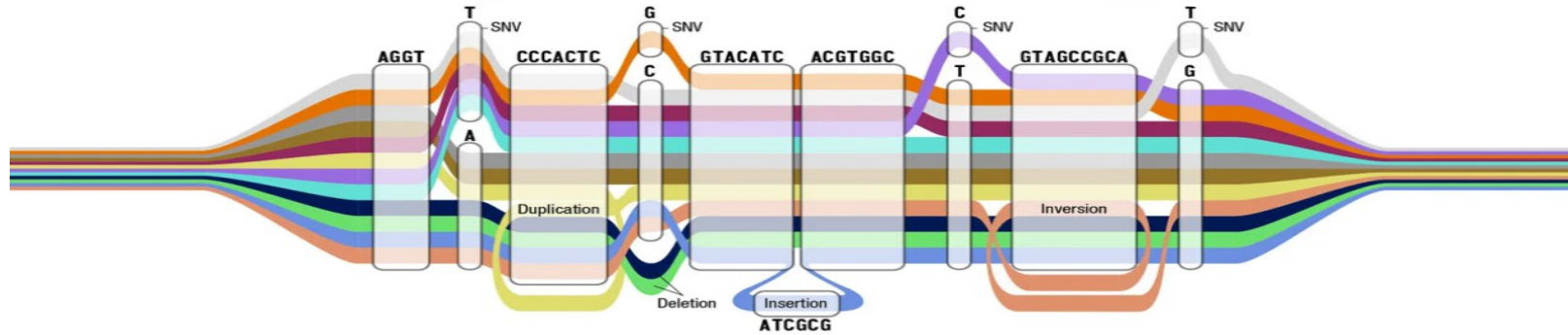
Nombre total de variants connus *

SNVs	786 500 648
InDels	122 583 462
SVs	1 199 117
Deletions	627 947
Duplications	258 882
CNVs	711
Insertions	296 184
Inversions	2 185
SV complexes	13 116



$>5 \times 10^6$ variants / individu

La signification clinique des variants



- Sans effet notable pour notre santé
- Expliquent nos différences



- Il a été démontré dans la littérature scientifique qu'ils causent le développement de maladies

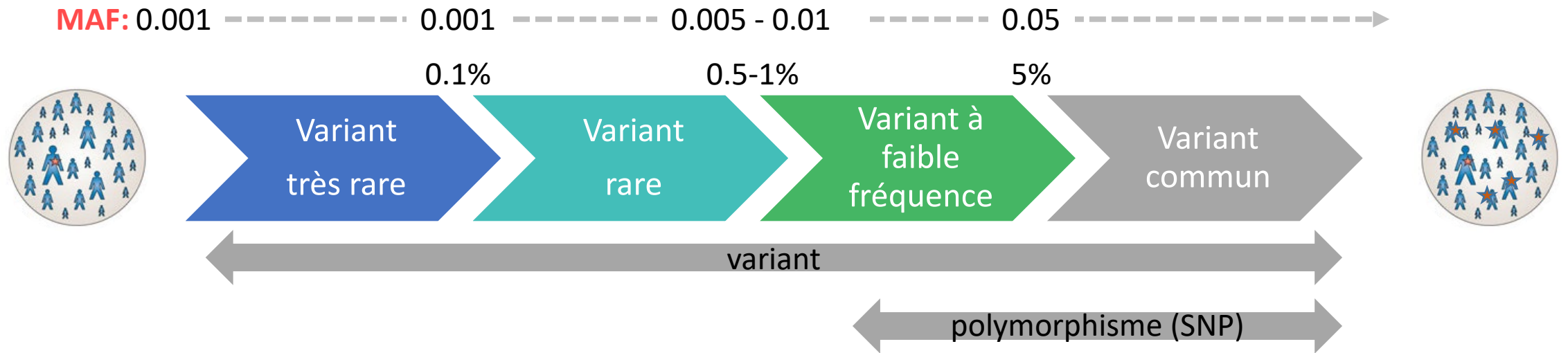
Plan du cours

A. La fréquence allélique du variant dans la population

B. L'impact fonctionnel des variants selon leur type

A. La fréquence allélique du variant dans la population

La fréquence allélique, du rare au fréquent





A l'échelle du génome humain, que pouvez-vous dire des variants communs ?

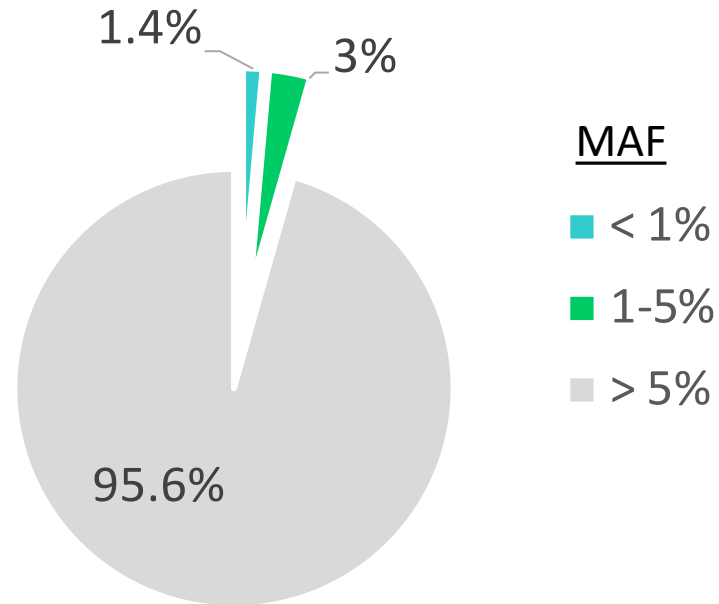


Figure 1: Répartition des variants du génome d'une personne en fonction de leur MAF (minor allele frequency).

Une réponse à sélectionner:

- A- les variants communs sont les plus nombreux de tous les variants d'un génome
- B- les variants communs ne sont pas les plus nombreux de tous les variants d'un génome
- C- je ne sais pas



Quelle fréquence allélique ne peut pas être considérée comme compatible avec une maladie génétique très rare ?

Une réponse à sélectionner:

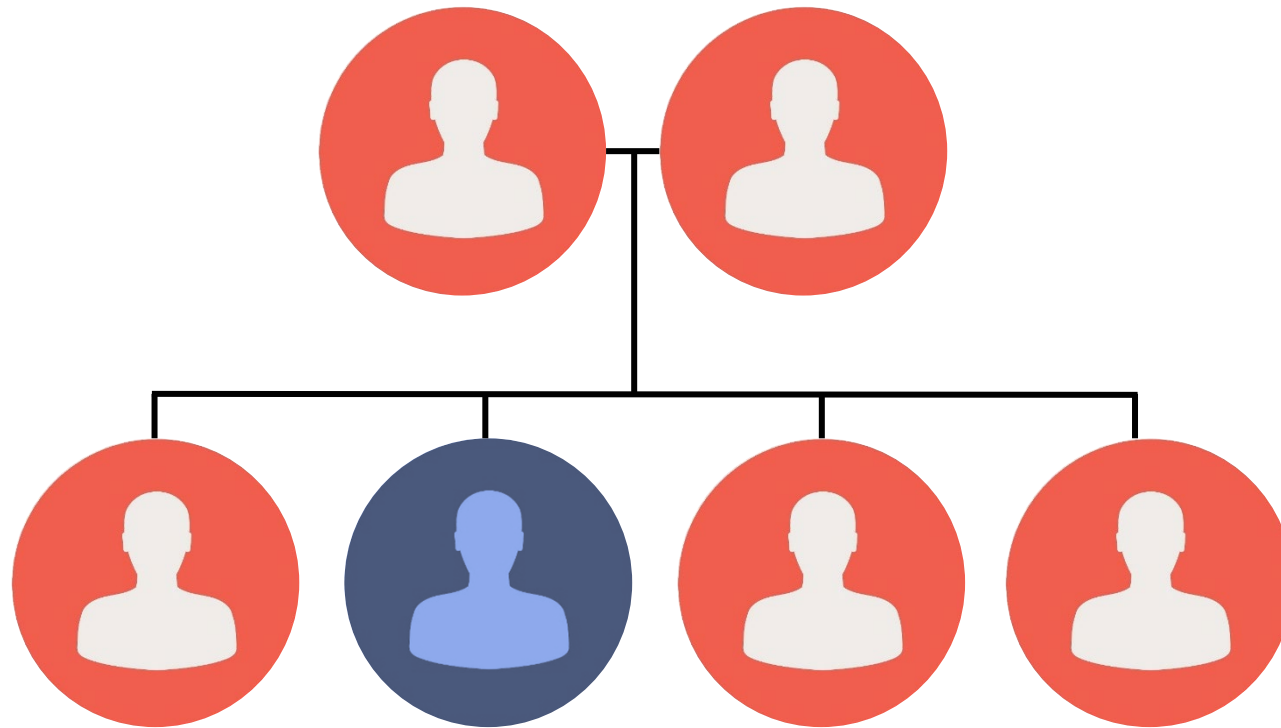
A- >5%

B- 1-5%

C- <1%

D- je ne sais pas

Tellement rare que je suis le/la première de la famille à développer cette maladie !

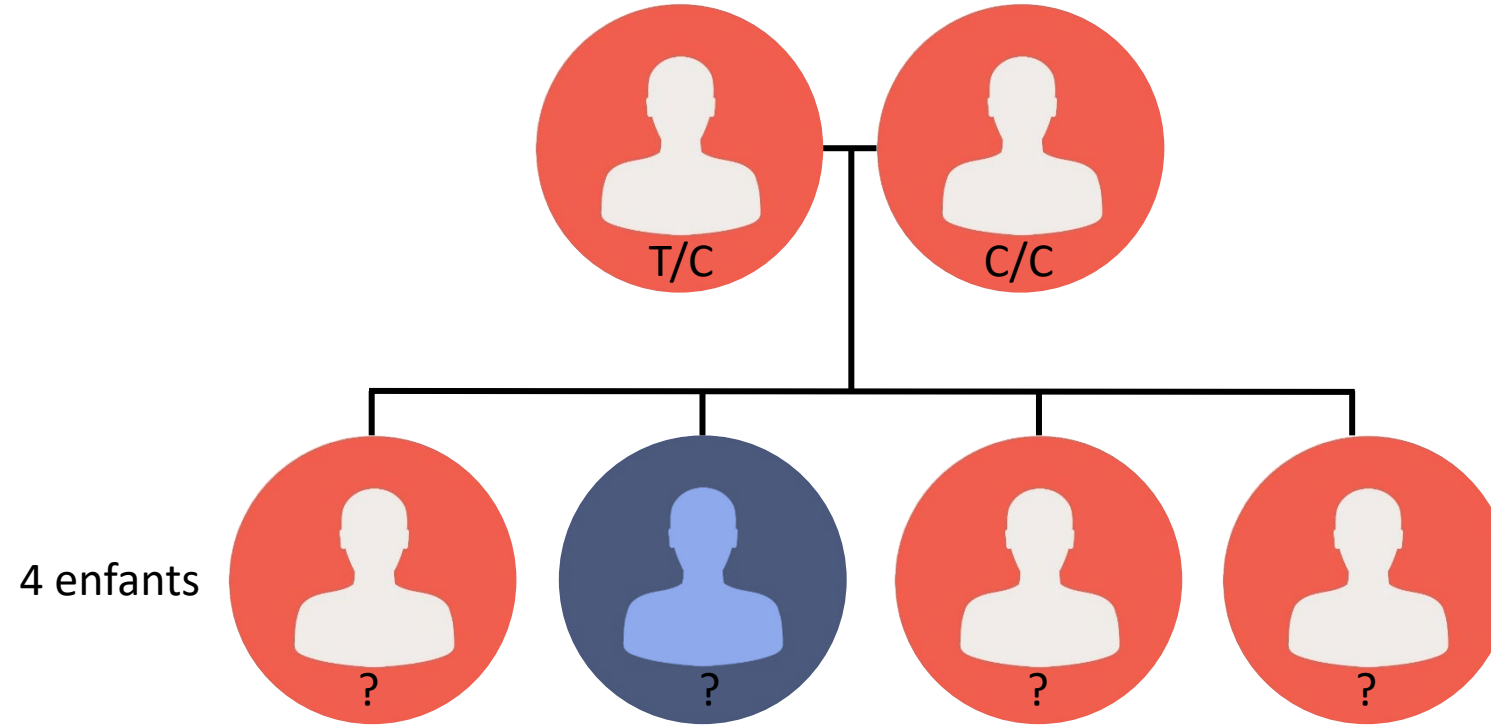




Séquence référence
.....CTGTATTCCT**T**GCCTGTCCAGGGAT.....

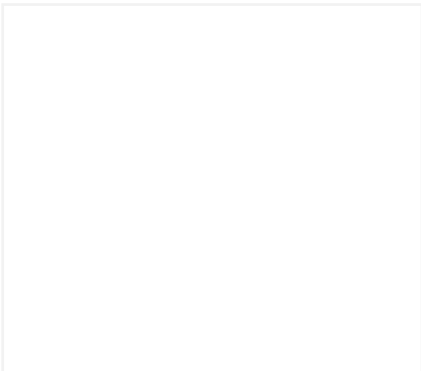
Séquence de la mère
Allèle mCTGTATTCCT**T**GCCTGTCCAGGGAT.....
Allèle pCTGTATTCCT**C**GCCTGTCCAGGGAT.....

Séquence du père
Allèle mCTGTATTCCT**C**GCCTGTCCAGGGAT.....
Allèle pCTGTATTCCT**C**GCCTGTCCAGGGAT.....

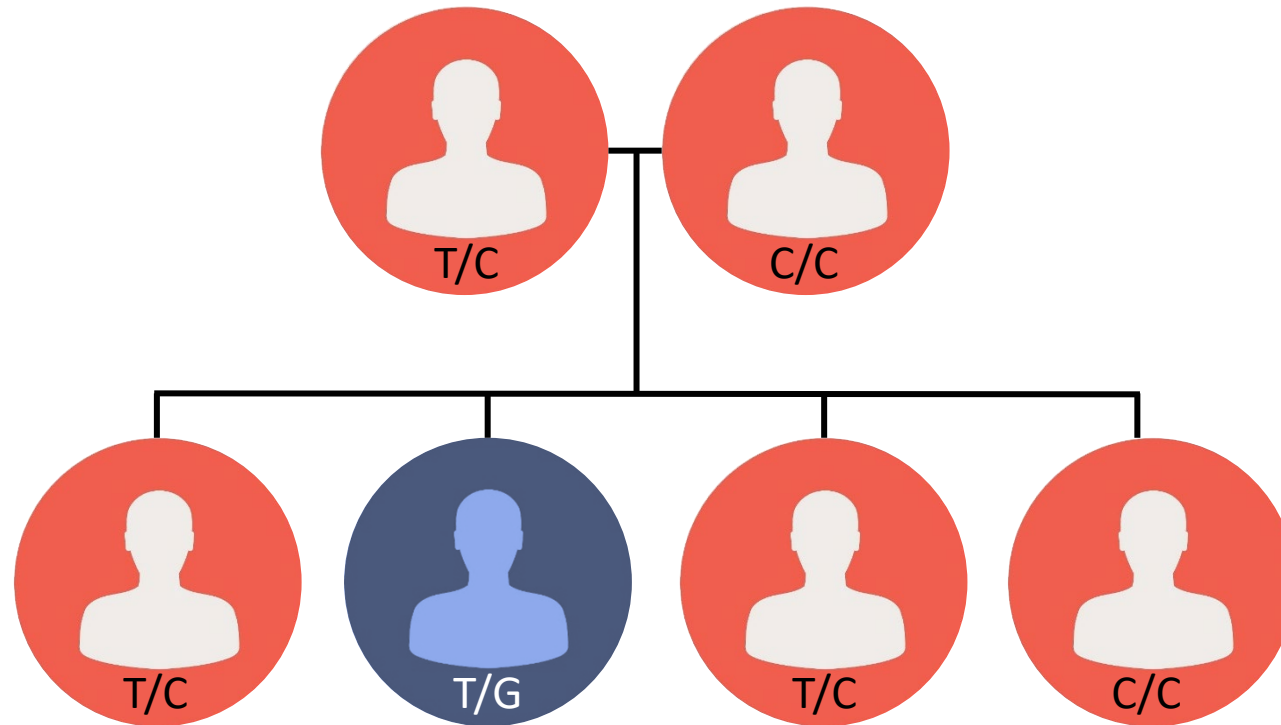


En théorie, quels seront les génotypes des enfants ?

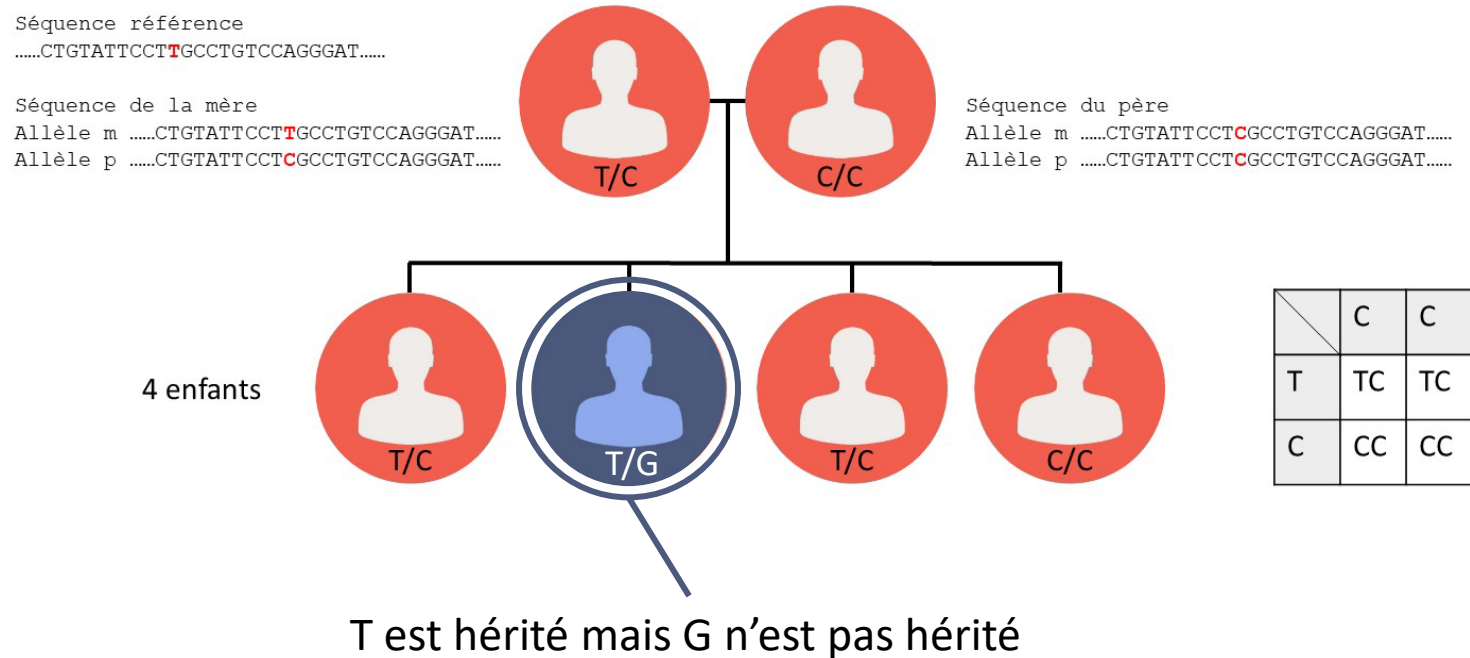
- A- TC
- B- CC
- C- TT
- D- je ne sais pas



En pratique, la situation suivante se produit mais très rarement ...



Les variants *de novo*



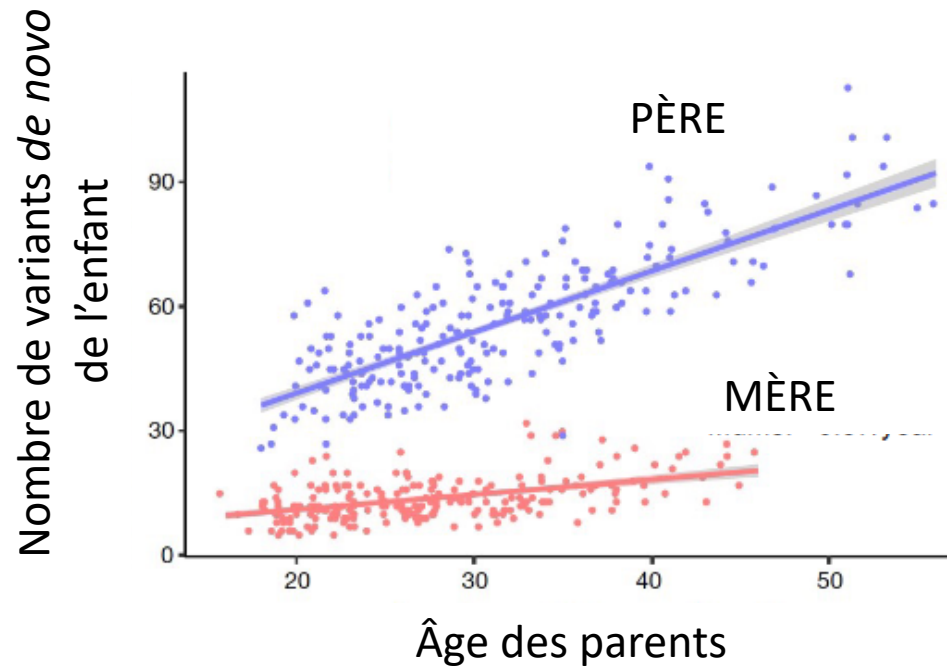
G est un variant *de novo* (variant spontané, non hérité):

- Ce variant est absent de l'ADN du sang des parents (non hérité).
- Ce variant pourra être transmis à la future descendance.
- **Ce variant apparaît lors d'une mitose durant la gamétogénèse.**
- **~70-100 variants *de novo*/individu.**
- 1 chance sur 100 millions qu'un nucléotide soit muté au hasard et à une position donnée dans le génome.

Variants *de novo*



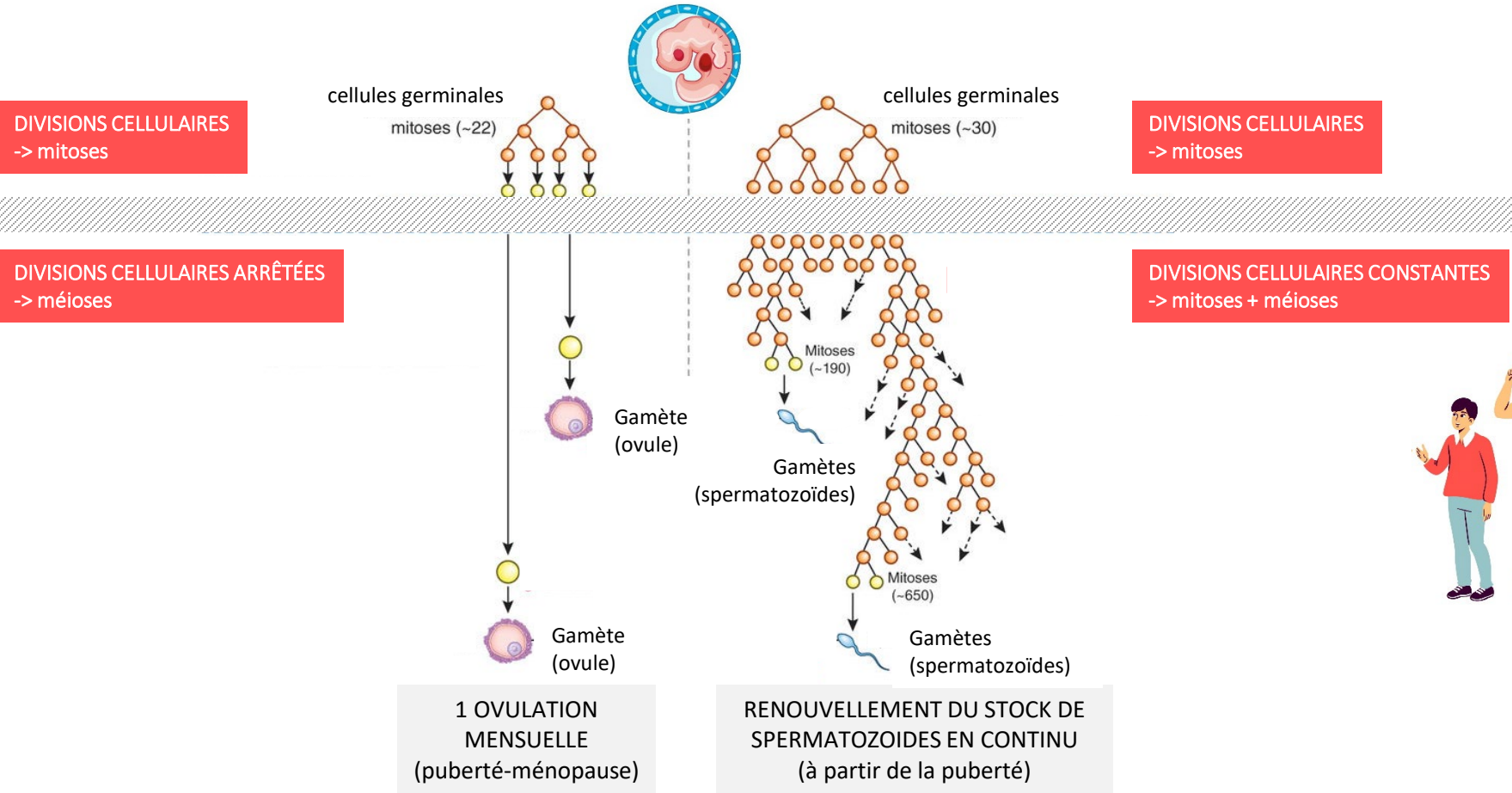
- $\frac{3}{4}$ des variants SNV *de novo* retrouvés dans le génome de l'enfant proviennent du génome des gamètes du père
- Observation: l'âge du père lors de la conception influence le nombre de variants *de novo* du génome de l'enfant



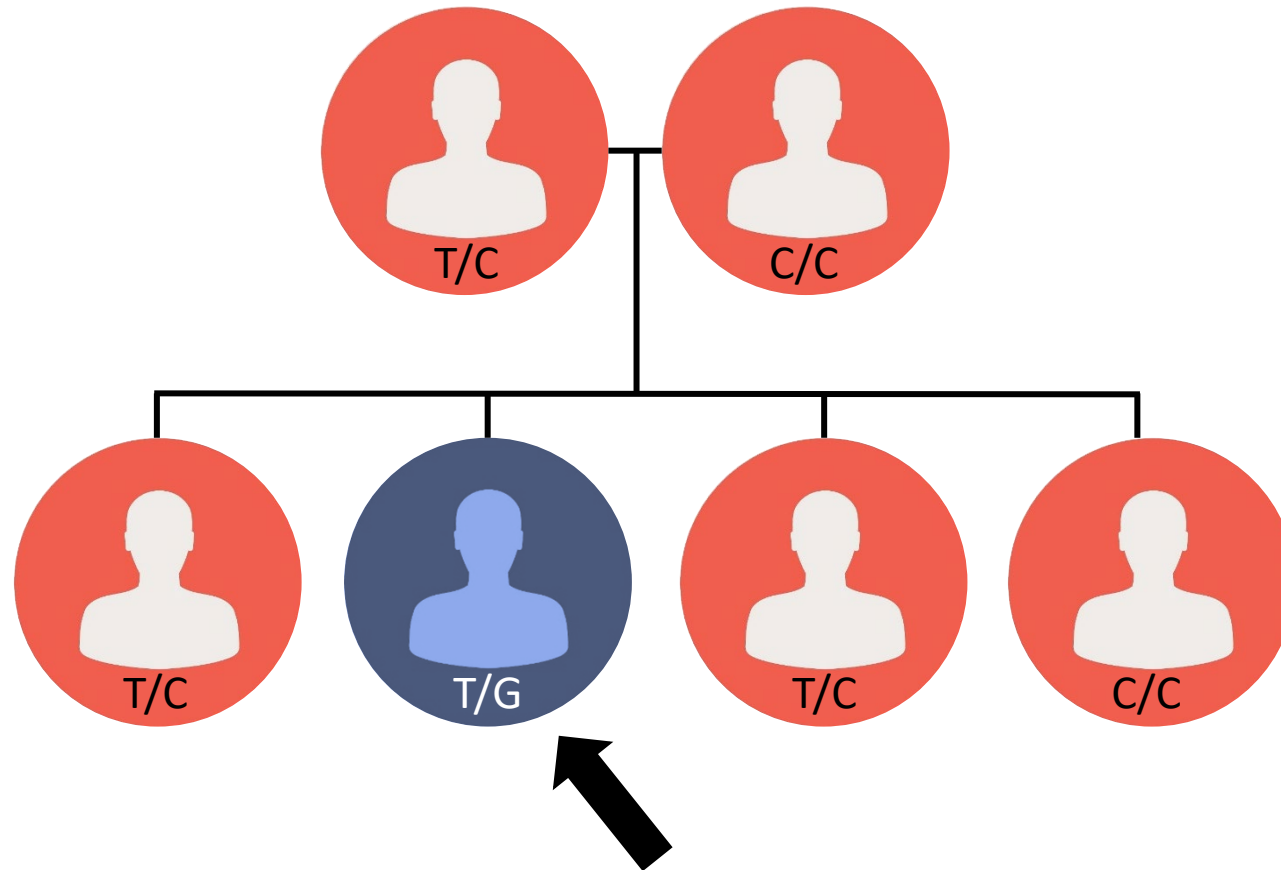
- Pourquoi ? L'hypothèse actuelle :
Parce que la fréquence d'apparition d'un variant *de novo* est corrélée au nombre de mitoses pendant la gamétogénèse

La gamétogénèse: ovogénèse versus spermatogénèse cf. cours S. Nef

enfance



Variant *de novo* et maladie rare



B. L'impact fonctionnel des variants selon leur type

1- L'impact fonctionnel des SNVs indels inversions (<50pb)

- 1a- Variants localisés dans des séquences **codantes**
- 1b- Altérations de l'épissage
- 1c- Variants localisés dans des séquences **non-codantes**

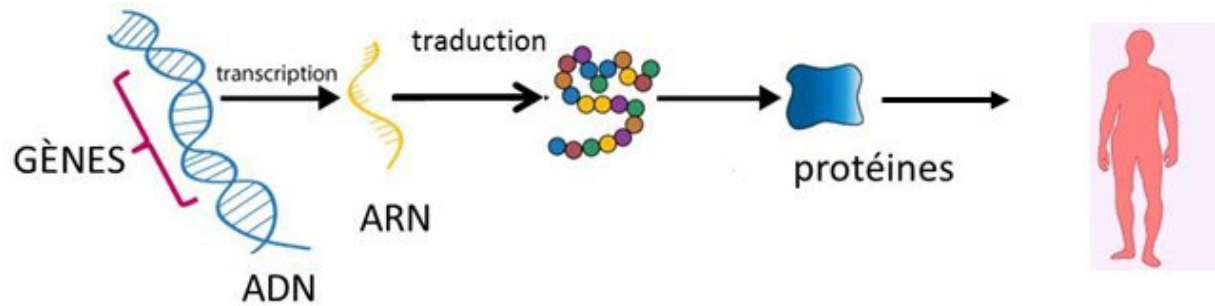


Est non-codant:

- A- Un intron
- B- Un exon
- C- Une séquence intergénique
- D- Un gène
- E- je ne sais pas

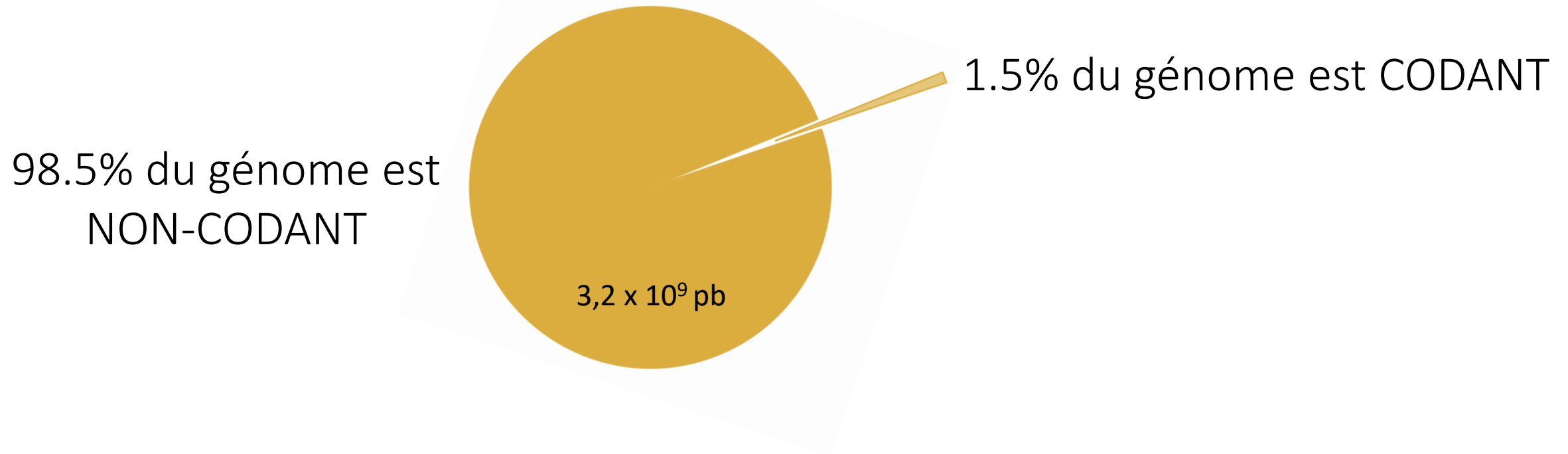


- Qu'est-ce qu'une séquence codante ?
Une séquence qui code pour une protéine



- Qu'est-ce qu'une séquence non-codante ?
Une séquence qui ne code pas pour une protéine

Génome humain: codant versus non-codant

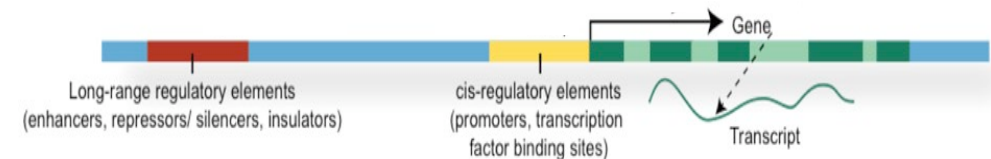
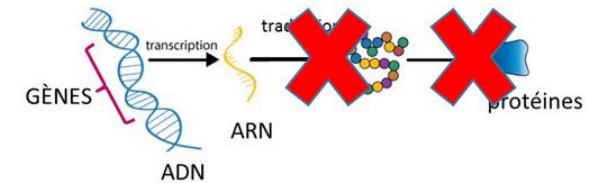
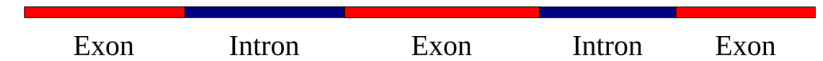




98.5% NON-CODANT

(i.e séquences du génome non-codantes pour des protéines)

- > 50% **séquences répétées**
- 25% du génome sont les **introns**
- 59 258 **gènes non-codants** pour des protéines
- **Séquences régulatrices** de la transcription
 - ~ 70 000 promoteurs
 - ~ 635 000 sites ADN de liaison à des facteurs de transcription
 - ~ 400 000 enhancers
 - etc.



1a- Variants localisés dans des séquences codantes



Le code génétique est redondant

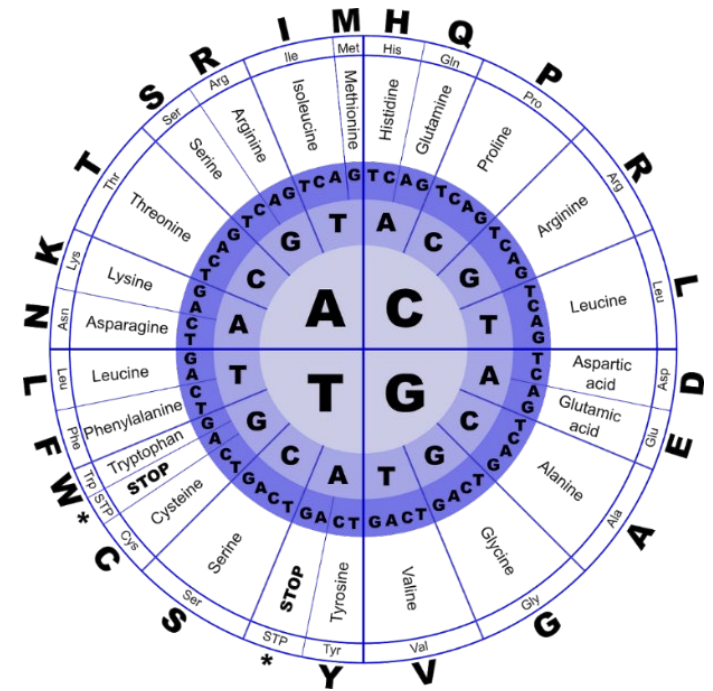
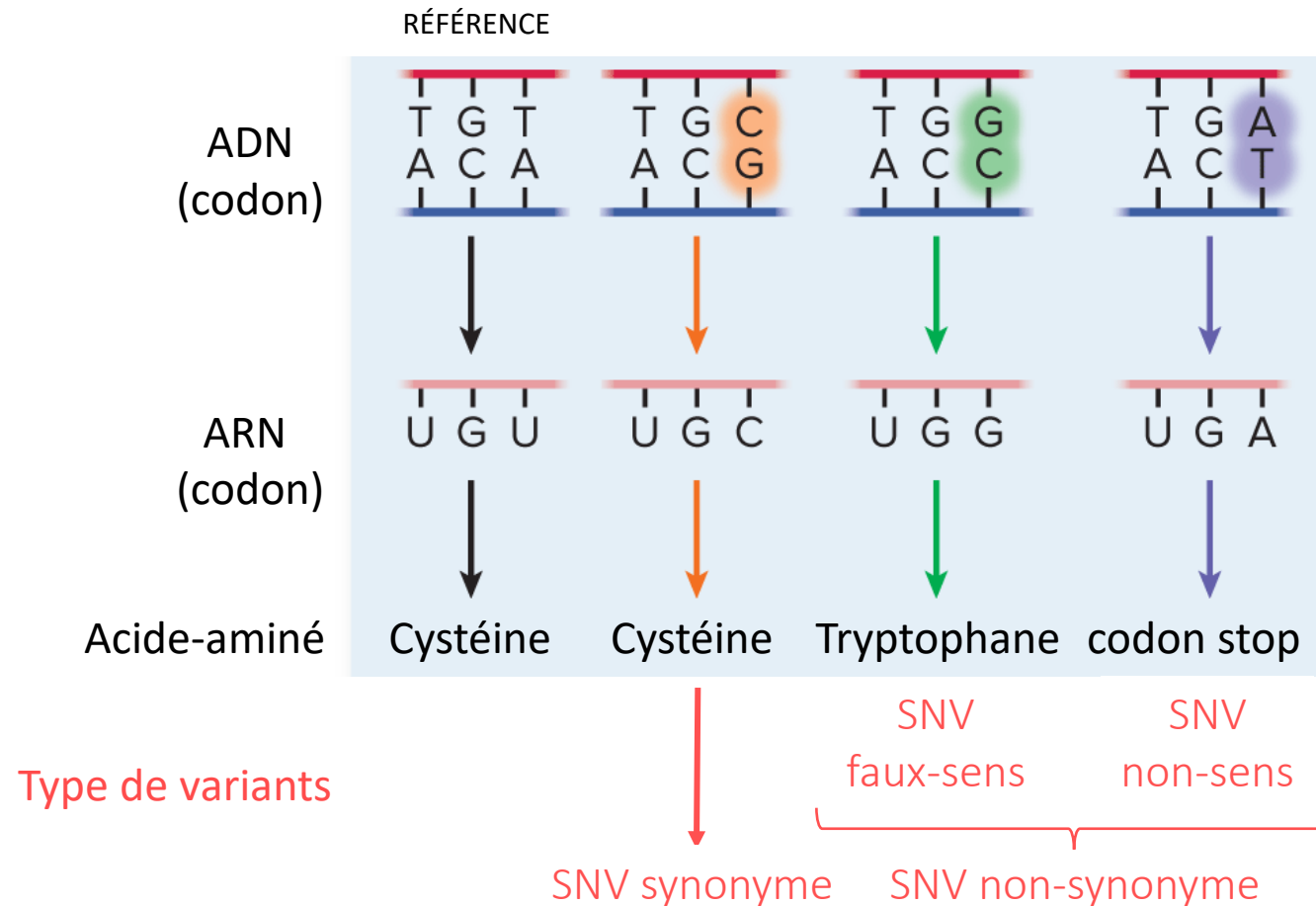
61 codons pour 20 acides-aminés et 3 codons stop

•Phénylalanine (Phe) :	UUU, UUC
•Leucine (Leu) :	UUA, UUG, CUU, CUC, CUA, CUG
•Isoleucine (Ile) :	AUU, AUC, AUA
•Méthionine (Met, Start) :	AUG
•Valine (Val) :	GUU, GUC, GUA, GUG
•Sérine (Ser) :	UCU, UCC, UCA, UCG, AGU, AGC
•Proline (Pro) :	CCU, CCC, CCA, CCG
•Thréonine (Thr) :	ACU, ACC, ACA, ACG
•Alanine (Ala) :	GCU, GCC, GCA, GCG
•Tyrosine (Tyr) :	UAU, UAC
•Histidine (His) :	CAU, CAC
•Glutamine (Gln) :	CAA, CAG
•Asparagine (Asn) :	AAU, AAC
•Lysine (Lys) :	AAA, AAG
•Aspartate (Asp) :	GAU, GAC
•Glutamate (Glu) :	GAA, GAG
•Cystéine (Cys) :	UGU, UGC
•Tryptophane (Trp) :	UGG
•Arginine (Arg) :	CGU, CGC, CGA, CGG, AGA, AGG
•Glycine (Gly) :	GGU, GGC, GGA, GGG
• Stop :	UAA, UAG, UGA

Les types de SNV et leurs impacts sur les acides aminés

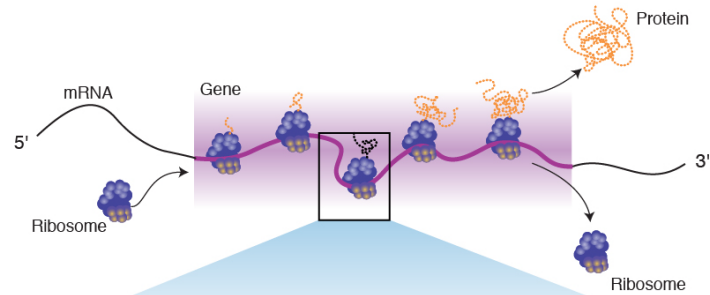
Les variants SNV synonymes n'ont pas d'impact connu car l'acide aminé est inchangé. Variant dite silencieux ou neutre.

Les variants SNV non-synonymes conduisent à un changement d'acide aminé.



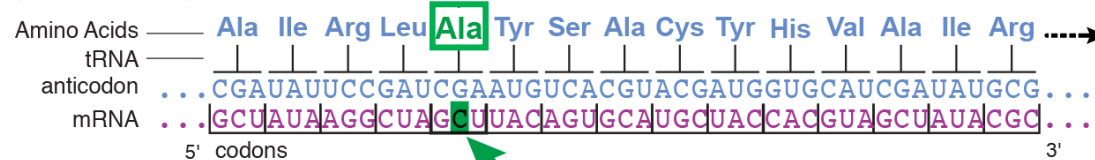
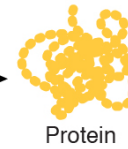
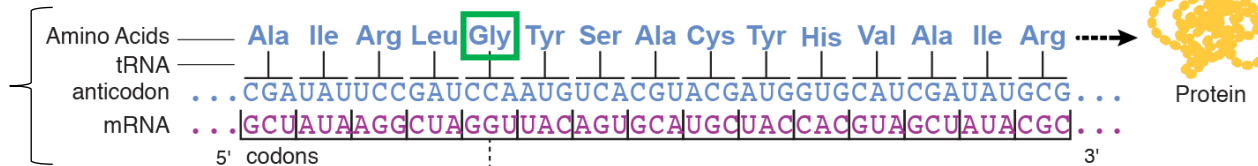
Les variants SNV faux-sens

(angl. missense SNV) engendrent un changement d'acide aminé.



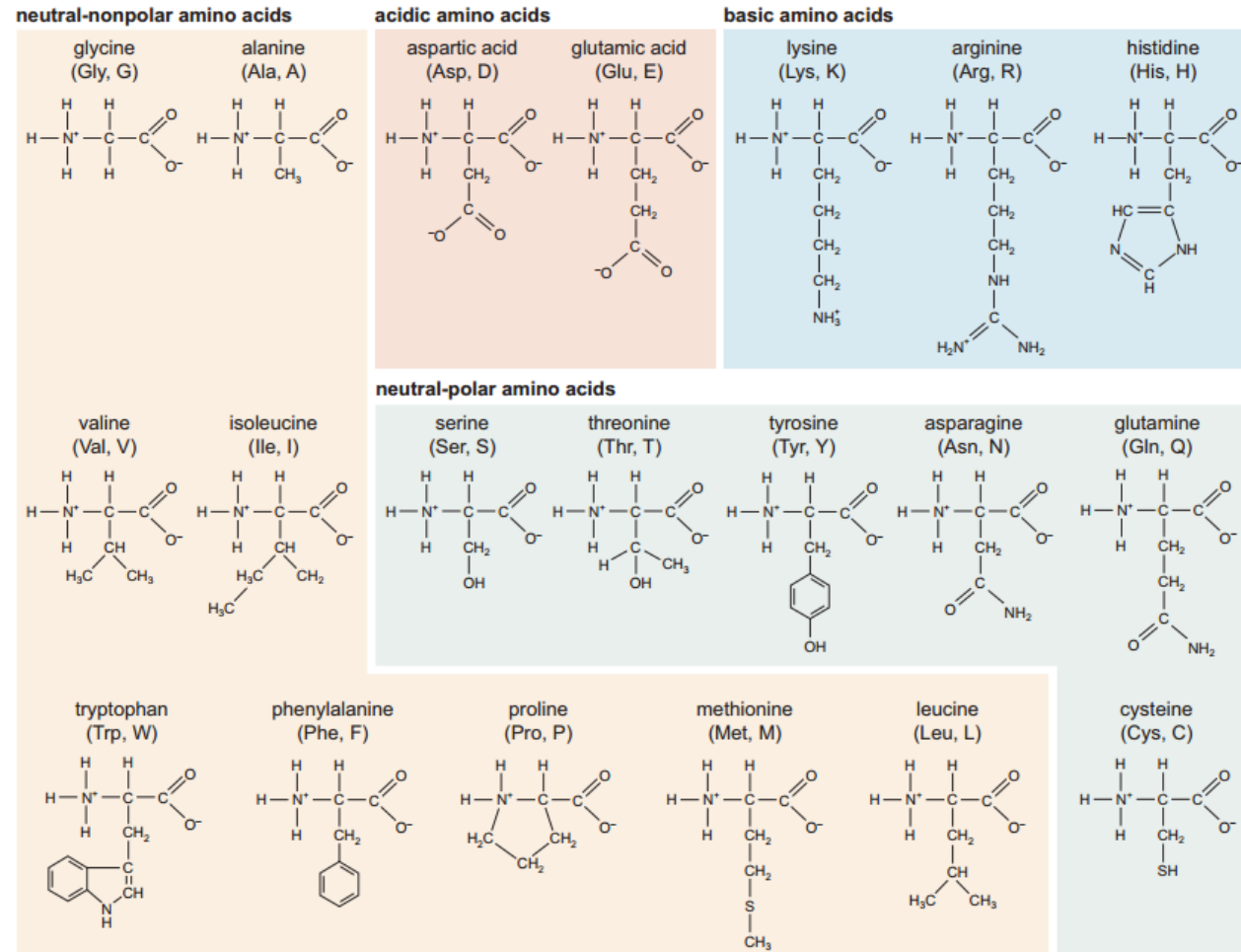
La séquence référence du génome humain ...GCUAUAAGGCUAGGUUACAGUGCAUGCAUACAGUAGCUAUACGC...

La séquence
référence
du génome
humain



SNV faux-sens

-> Impact séquence peptidique : Glycine remplacée par Alanine



cas 1 (conservative):

le codon spécifie un acide aminé biochimiquement équivalent. ex: Lys -> Arg

cas 2 (non conservative):

Le codon spécifie un acide aminé biochimiquement différent. ex: Lys -> Thr



Séquence référence:

... ACC GAC TAT ATA TAT CCG CAC TAC TTC GAC ACT

Séquence avec un variant:

... ACC GAC TAC ATA TAT CCG CAC TAC TTC GAC ACT

Quel type de variant ?

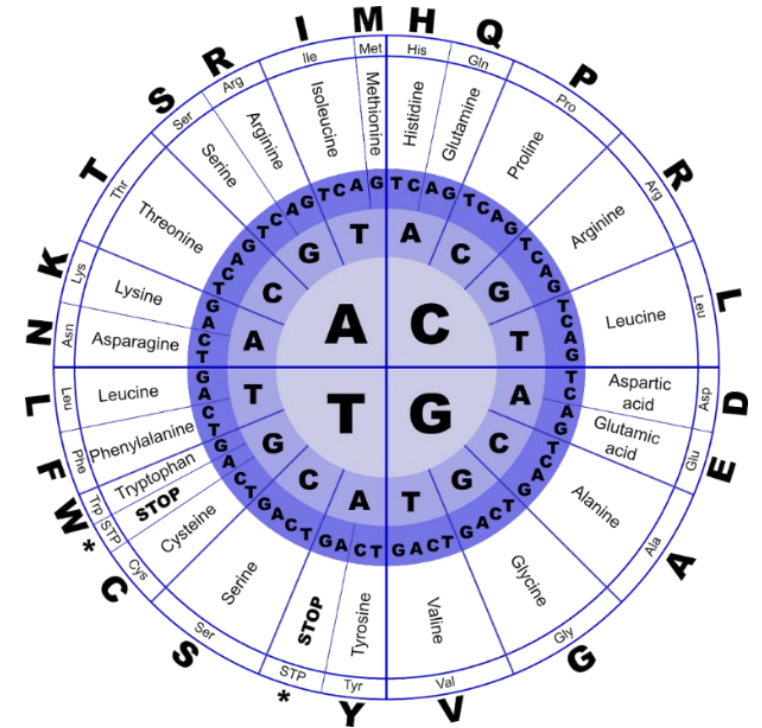
A- SNV synonyme

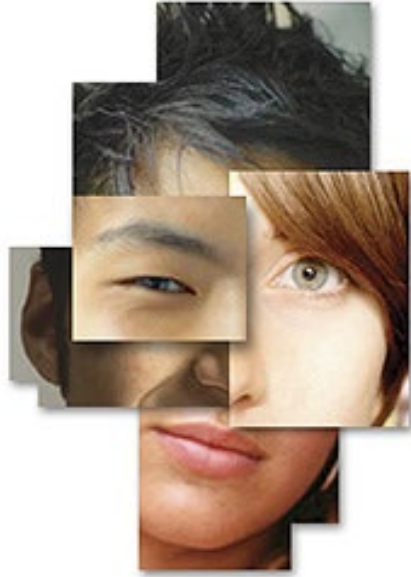
B- SNV faux-sens

C- SNV non-sens

D- indel synonyme

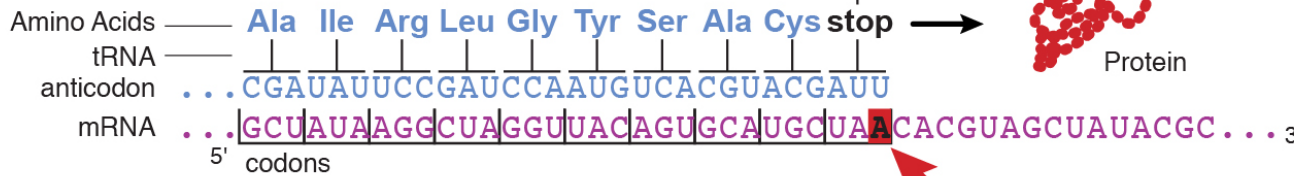
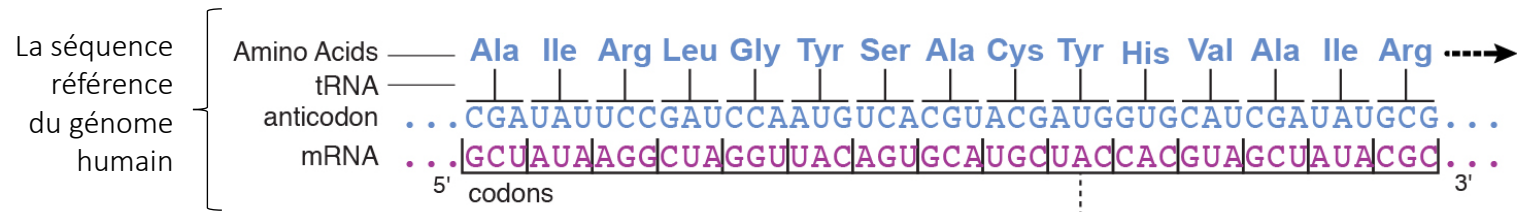
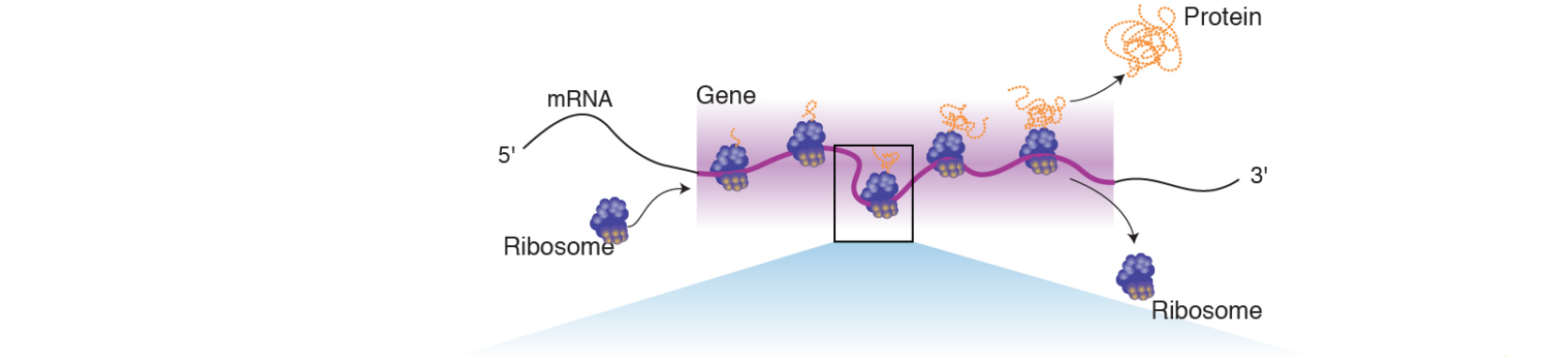
E- je ne sais pas





Les variants SNV non-sens

(angl. nonsense SNV) engendrent une protéine tronquée (plus courte) à cause d'une interruption prématurée de la traduction, le plus souvent rendant la protéine non-fonctionnelle (perte de fonction).



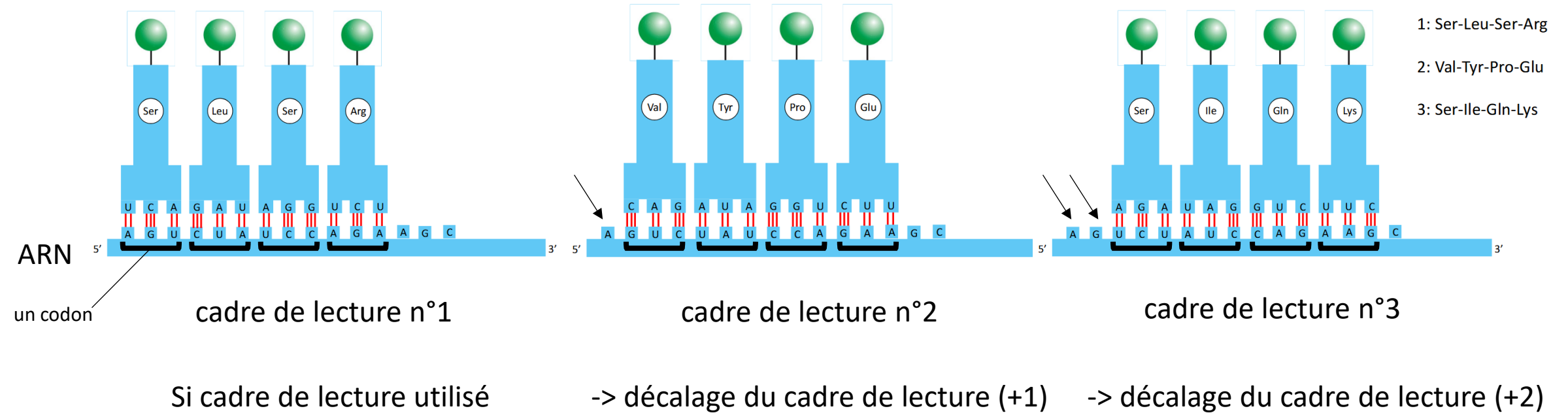
SNV non-sens

-> apparition d'un codon stop prématuré

Codon ADN STOP:
TAA
TAG
TGA

LES TROIS CADRES DE LECTURE POSSIBLES de 5' en 3'

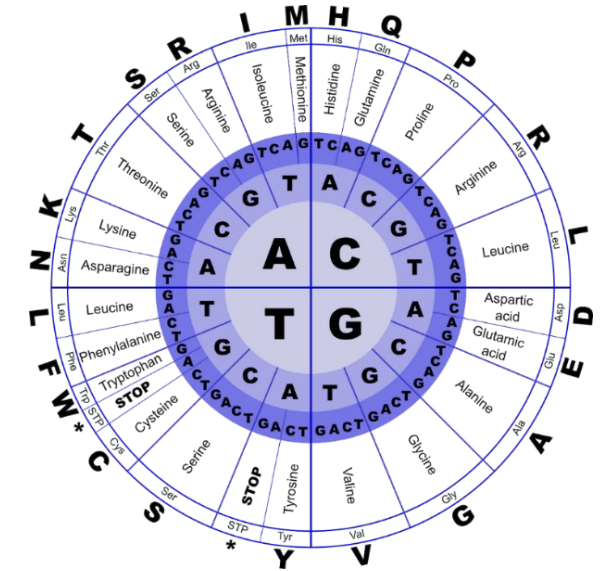
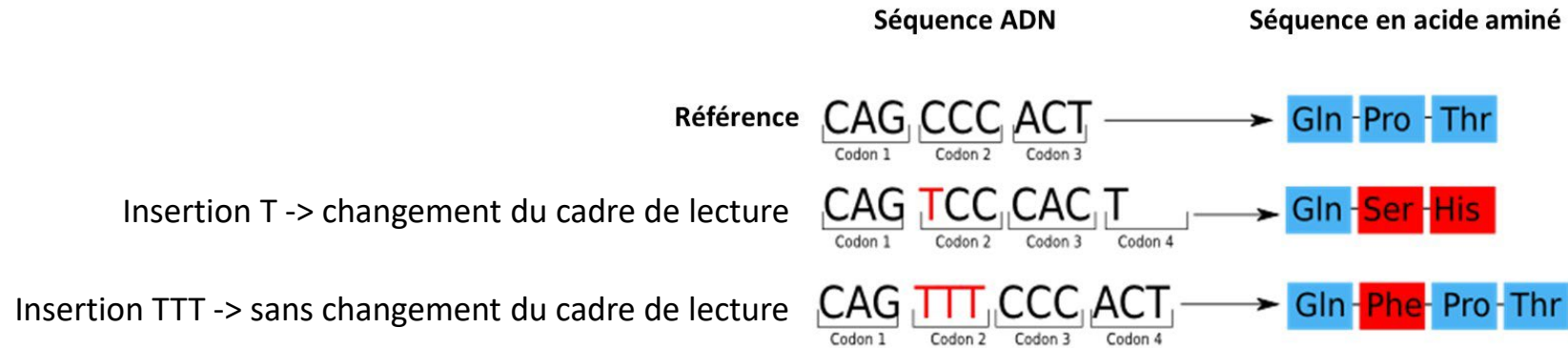
COURS du prof. Oliver Hartley



Impact des indel sur les acides aminés



(angl., in frame (sans décalage), frameshift (avec décalage))

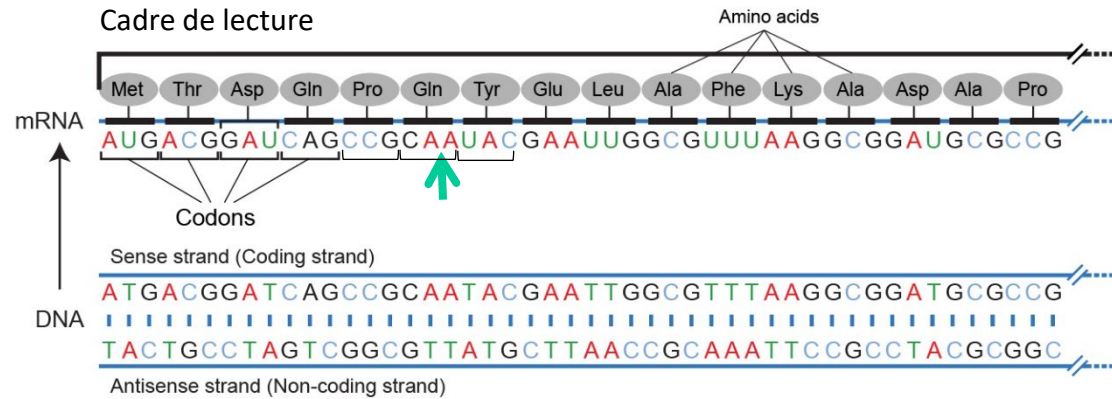


Questions à se poser pour un indel : Induit-il un décalage du cadre de lecture ?

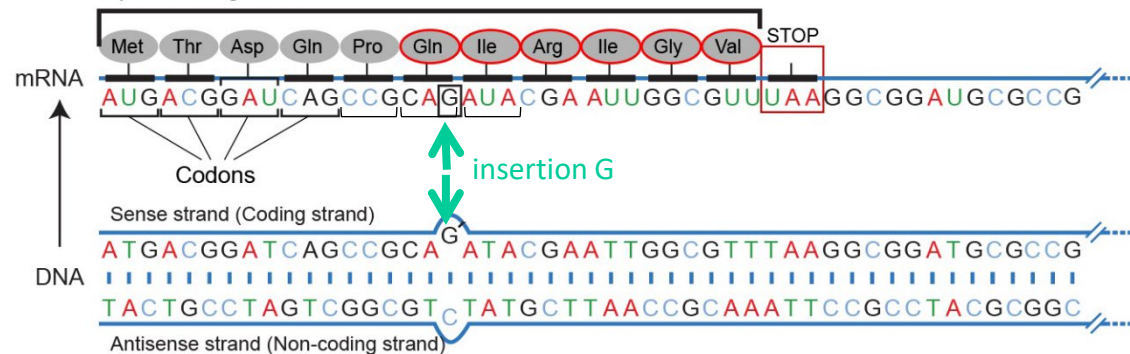
-> Est-il un multiple de 3 ?

-> Analyse de la conséquence sur les acides aminés

- La séquence référence du génome humain



- EXEMPLE: un insertion conduisant à un décalage du cadre de lecture (+1), changement de la composition en acide-aminé et introduction d'un codon stop prématuré



Impact des inversions

Séquence référence du génome humain

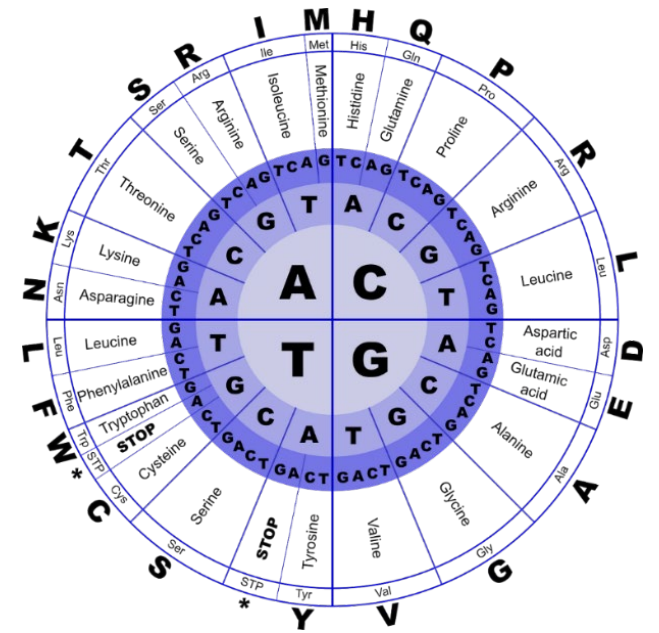
... ATTGGCCTTAACCCCGATTATCAGGA ...
I G L N P R L S G

Ex: une inversion de 7 paires de base

... ATTGGCCTTAAC**TAGCCCC**TATCAGGA ...
I G L N **STOP**

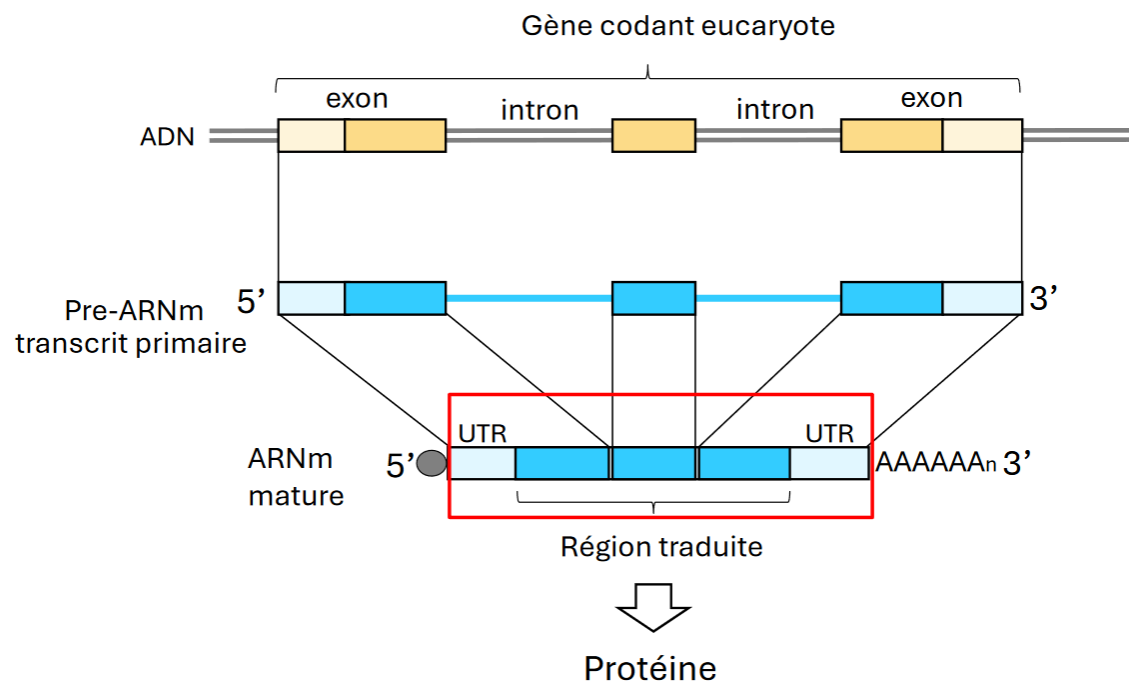
Ex: une inversion de 7 paires de base

... ATTGGCC**CAATT**CCCGATTATCAGGA ...
I G **P I** P R L S G



1b- Altérations de l'épissage

3. Élimination des introns = épissage (splicing)

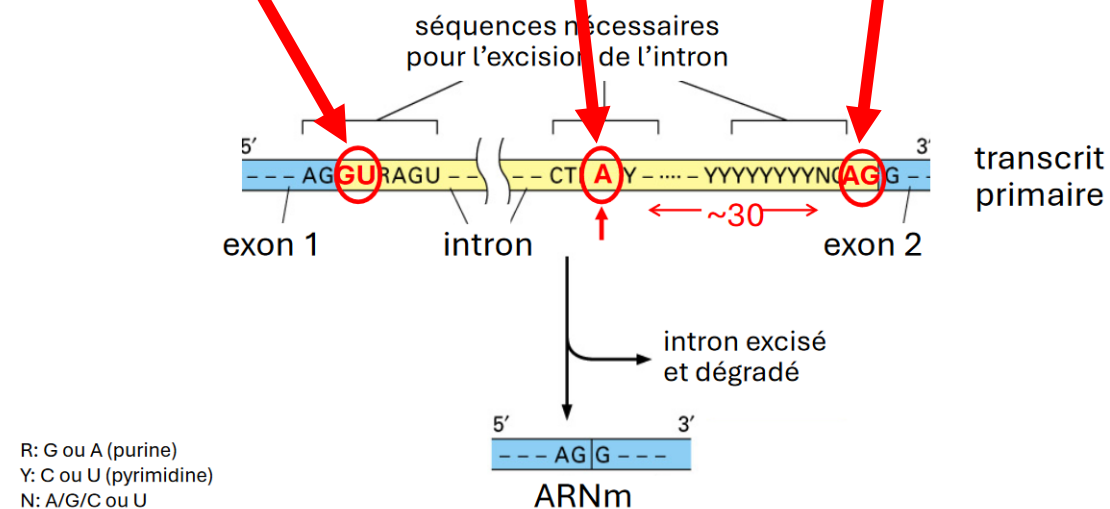


Site donneur

Site accepteur

Point de branchement

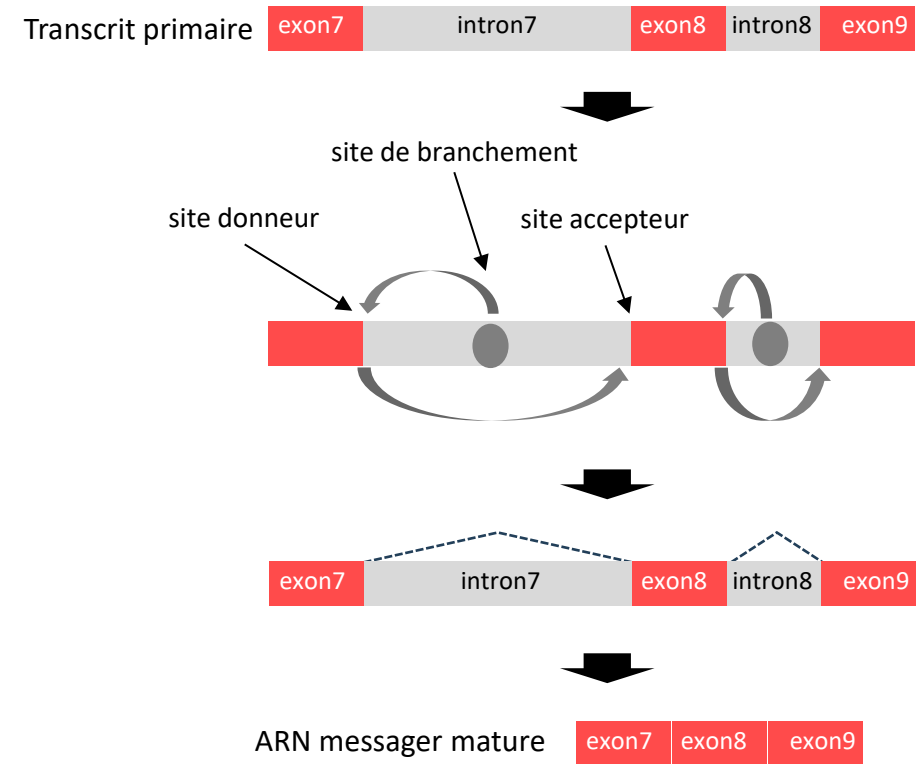
Les signaux d'épissage sur l'ARNm



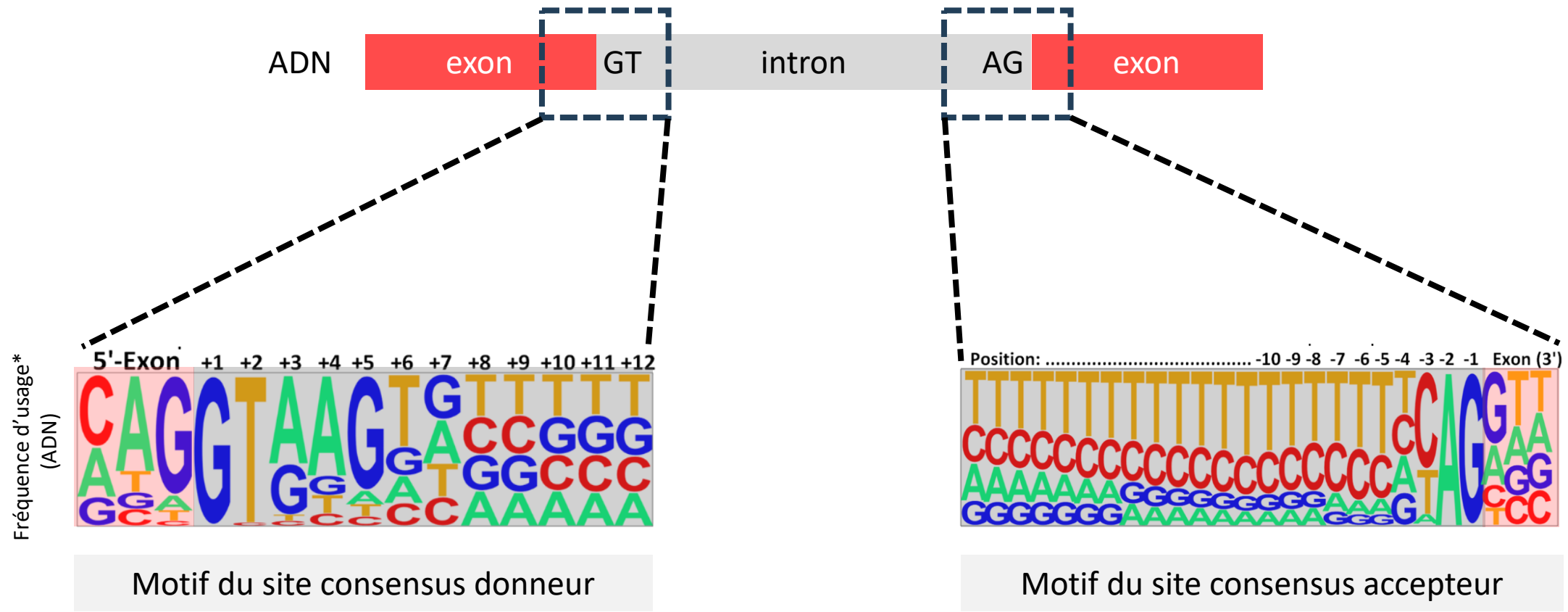
L'épissage

L'épissage en quelques points clefs pour ce cours :

- initiation par la reconnaissance du point de branchement
- identification du site d'épissage 5' (site donneur)
- identification du site d'épissage 3' (site accepteur)
- élimination de l'intron entre les 2 exons



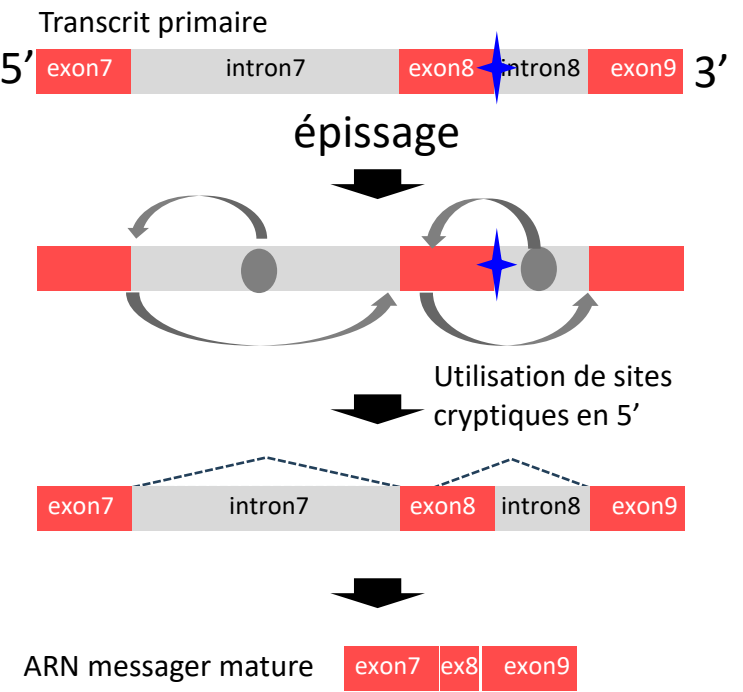
Les sites donneurs et accepteurs d'épissage



* La hauteur de la lettre indique sa fréquence d'usage : plus elle est haute, plus la base est utilisée et donc importante dans la reconnaissance du motif

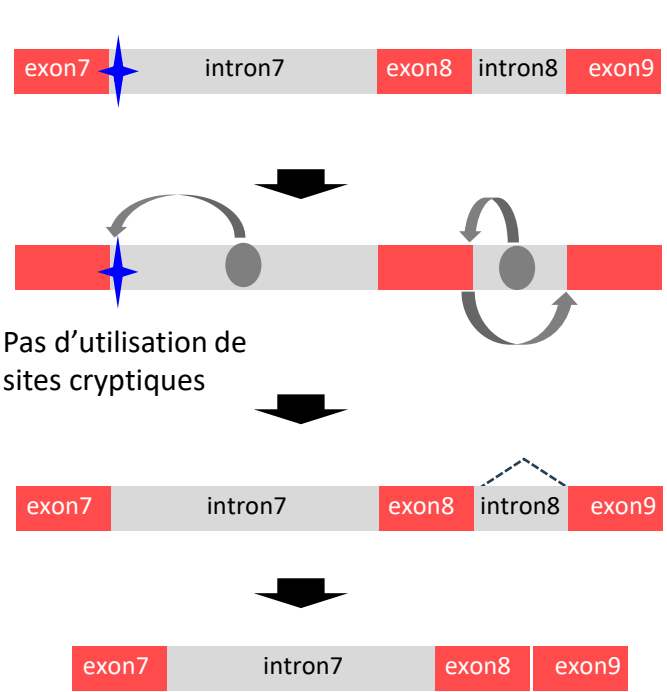
Exemples de défauts d'épissage causés par des variants*

Modification d'un exon



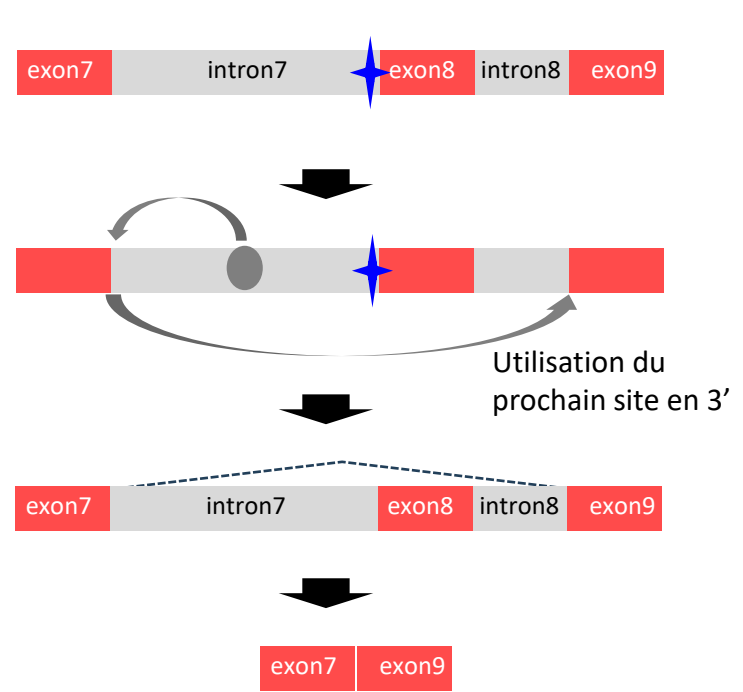
➔ Modification d'un exon et modification possible du cadre de lecture

Rétention d'intron



➔ Rétention d'intron et modification possible du cadre de lecture

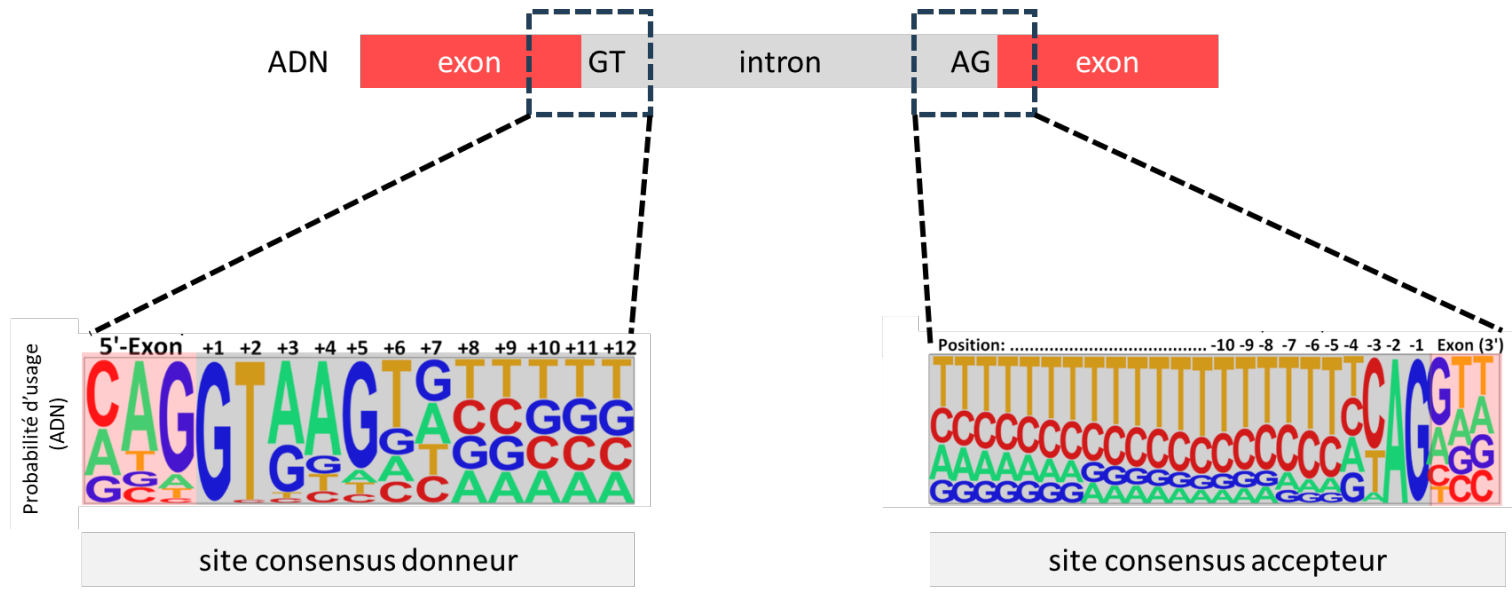
Saut d'exon



➔ Elimination d'un exon et modification possible du cadre de lecture

Les sites donneurs et accepteurs d'épissage

- Les maladies génétiques causées par un défaut d'épissage démontrent l'importance des bases suivantes :



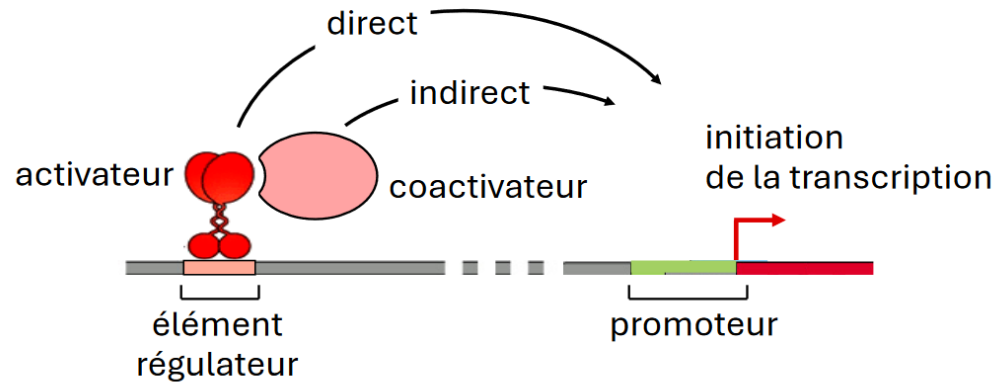
+1 (G) du site donneur
 +2 (T) du site donneur
 +5 (G) du site donneur

-1 (G) du site accepteur
 -2 (A) du site accepteur

1c- Variants localisés dans des séquences non-codantes

Dans les exemples suivants, un variant localisé dans une séquence non-codante peut modifier le niveau de transcription d'un gène selon le type cellulaire.

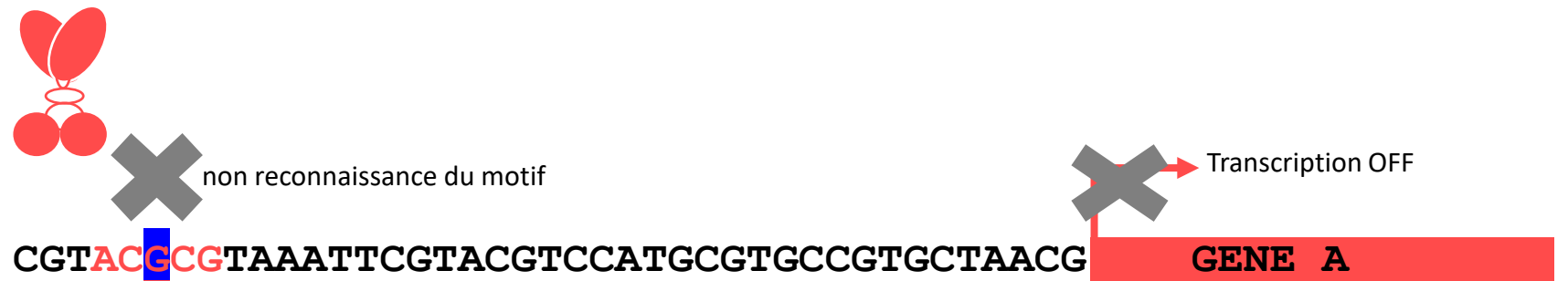
Cours 3 de la prof. Martina Valentini



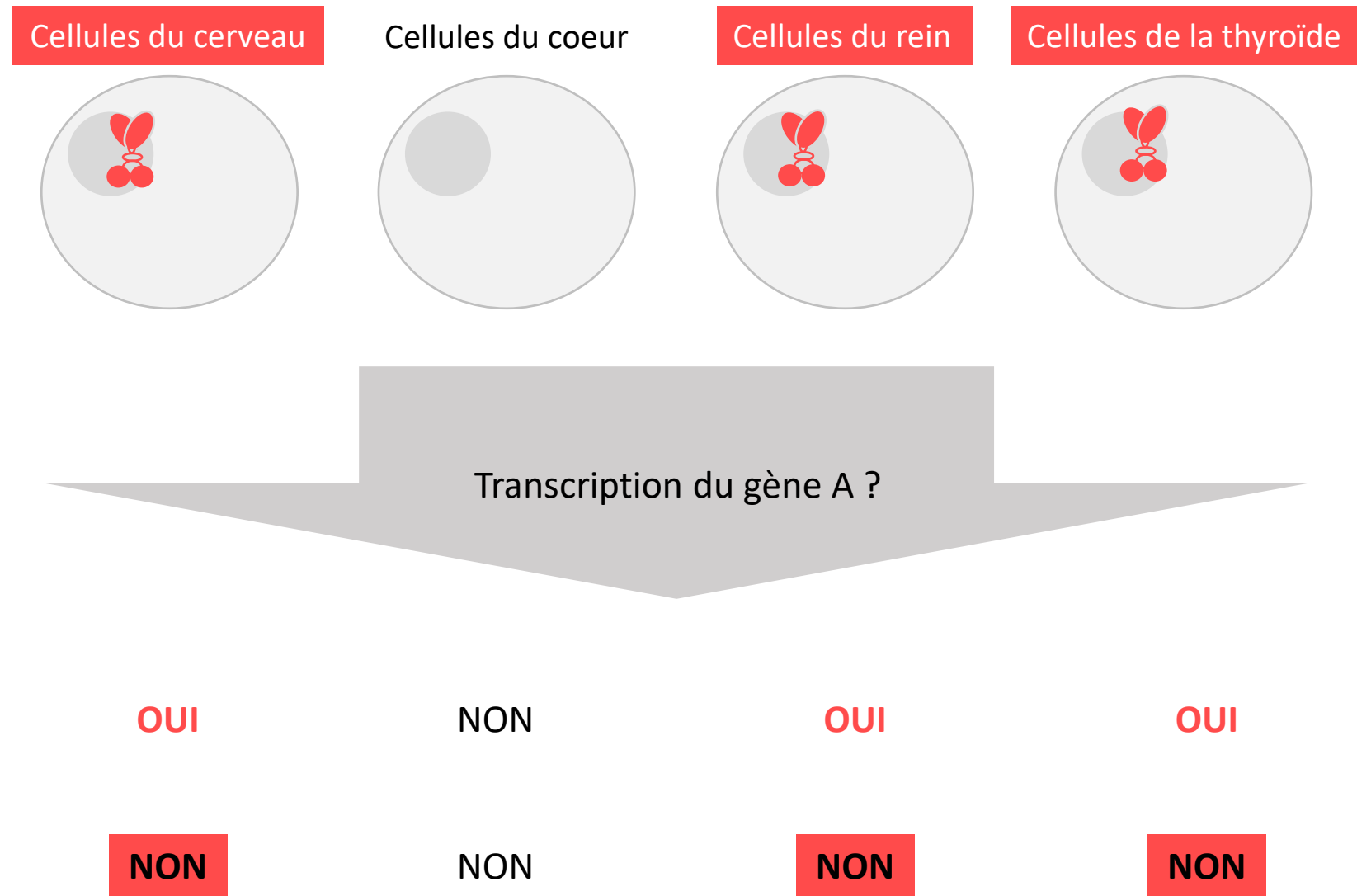
Séquence référence:



Séquence avec un SNV T>G :



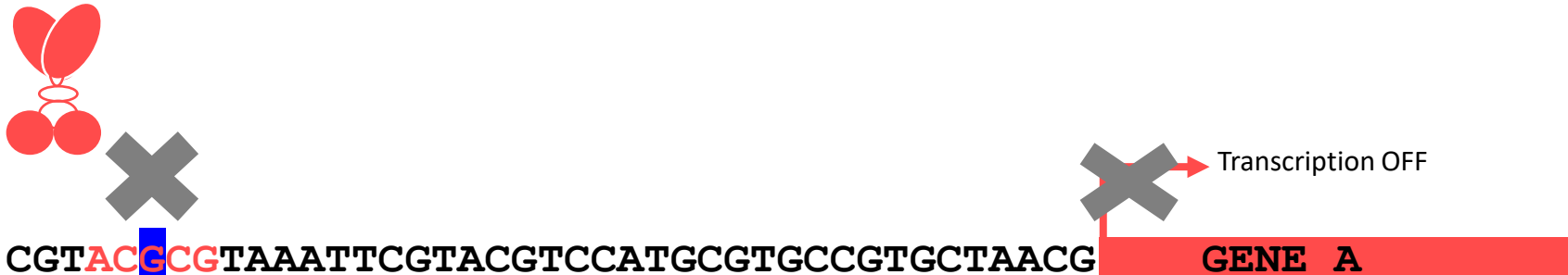
L'activateur est présent que dans certains tissus/cellules:



Séquence référence:



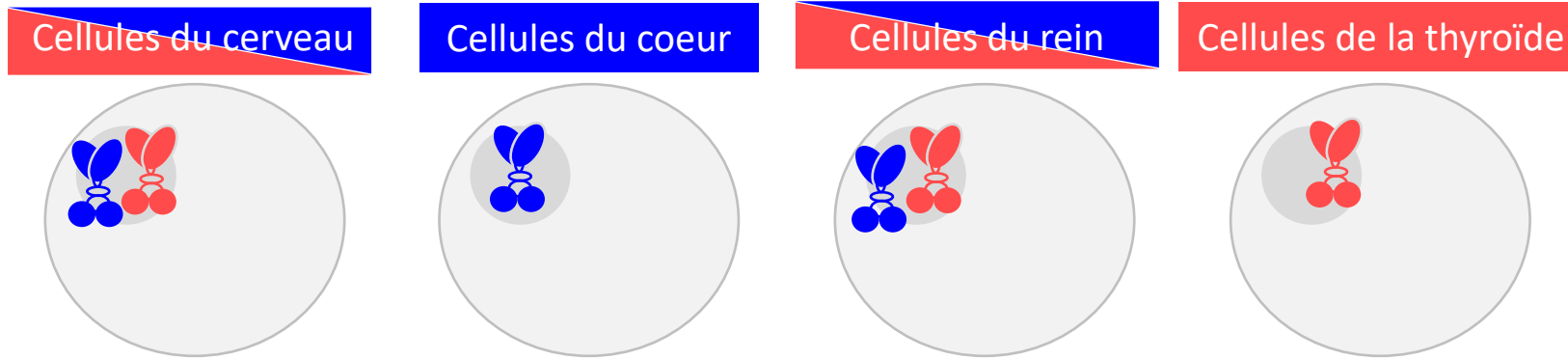
Séquence avec un SNV T>G :



Séquence avec un SNV T>G :



Les activateurs sont présents dans certains tissus/cellules :



Transcription du gène A ?

Séquence référence

OUI

NON

OUI

OUI

Séquence SNV T>G

OUI

OUI

OUI

NON

2- Autres

2a- Les impacts des CNV

2b- L'impact des anomalies numériques de chromosomes

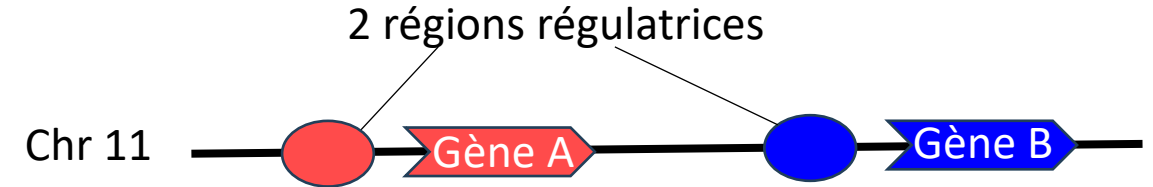
2c- L'impact des translocations

2a- Les impacts des CNV

Analyser la position des CNV par rapport aux gènes

Exemple de délétions

L'annotation de la séquence référence du génome humain est:



-> Elimine une région régulatrice



-> Elimine un gène, une région régulatrice et rapproche une autre région régulatrice

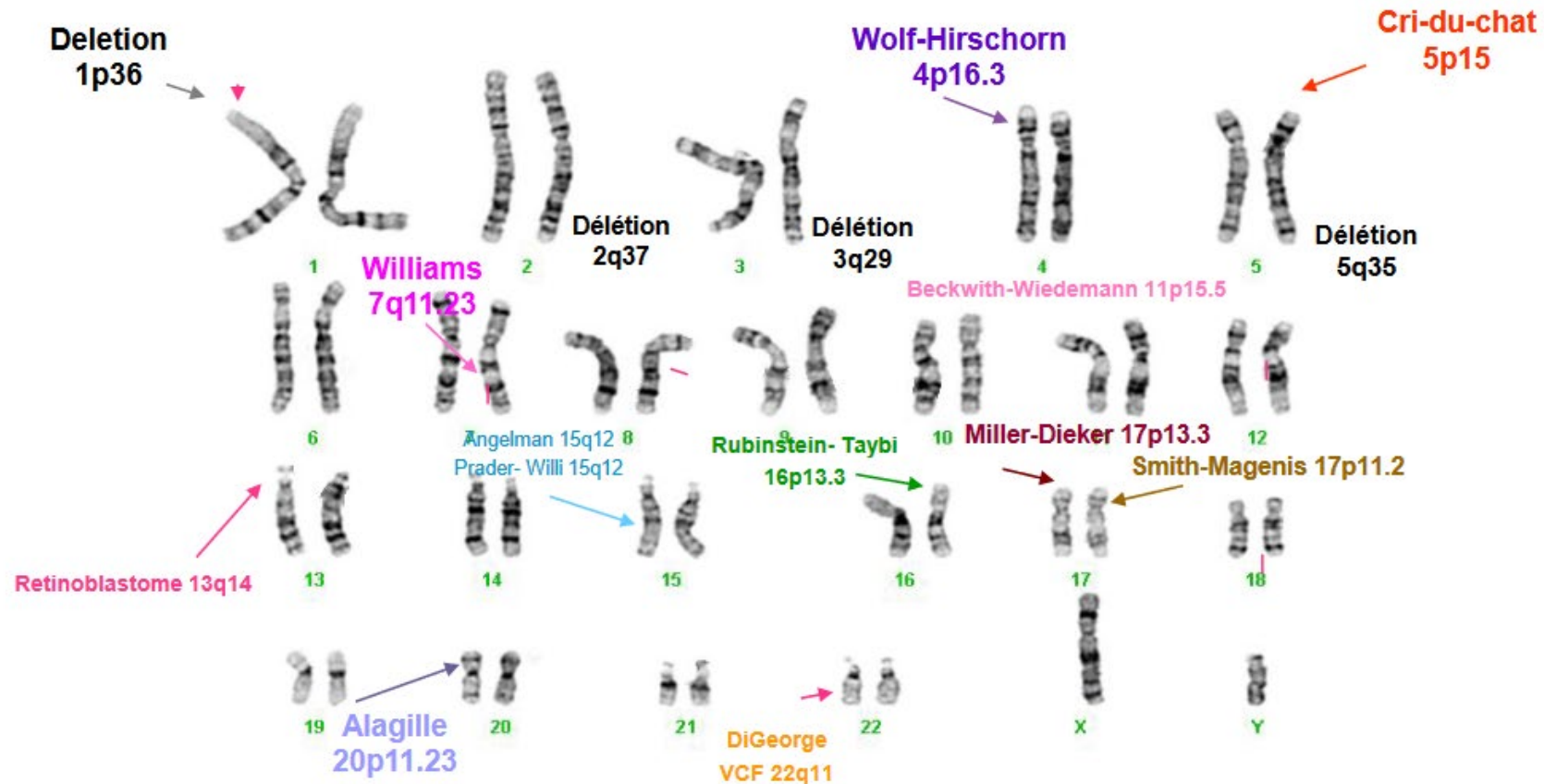


-> Fusionne 2 gènes

Les syndromes causés par des SV



- Ici délétion ou gain de copies de régions génomiques (CNV)
- >100 maladies rares
- Symptômes cliniques hétérogènes car plusieurs organes touchés

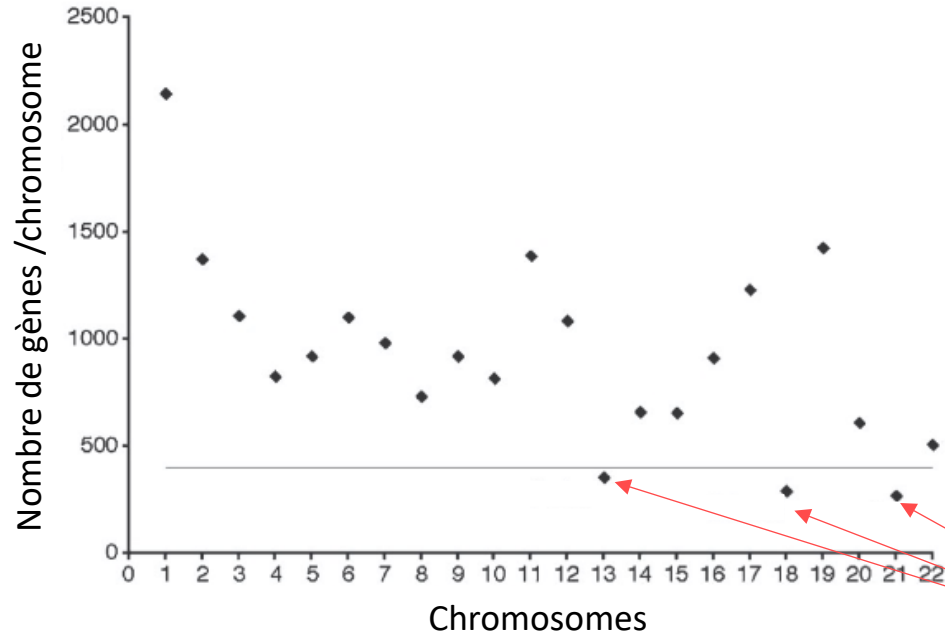


2b- L'impact des anomalies numériques de chromosomes



Les Monosomies

- Monosomie complète d'un autosome -> létale *in utéro*
- Monosomie du Y -> létales *in utéro* (45,Y0)
- Monosomie du X -> Syndrome de Turner (45,X0)



Les Trisomies des autosomes

- Les Trisomies 21, 13 et 18 sont les seules trisomies autosomiques humaines pas létales *in utéro*

2c- L'impact des translocations

Translocation équilibrée versus déséquilibrée



Type de translocation	Perte/gain d'ADN ?	Effet sur le porteur de la translocation <i>de novo</i>	Risque pour la descendance (cf. TD de génétique humaine)
Équilibrée	Non*	Pas manifestation clinique connue	Modéré
Déséquilibrée	Oui	Souvent Maladie ou Syndrome	Élevé

*la perte du bras court d'un chromosome acrocentrique n'est pas considéré comme une perte car l'information est redondante dans le génome

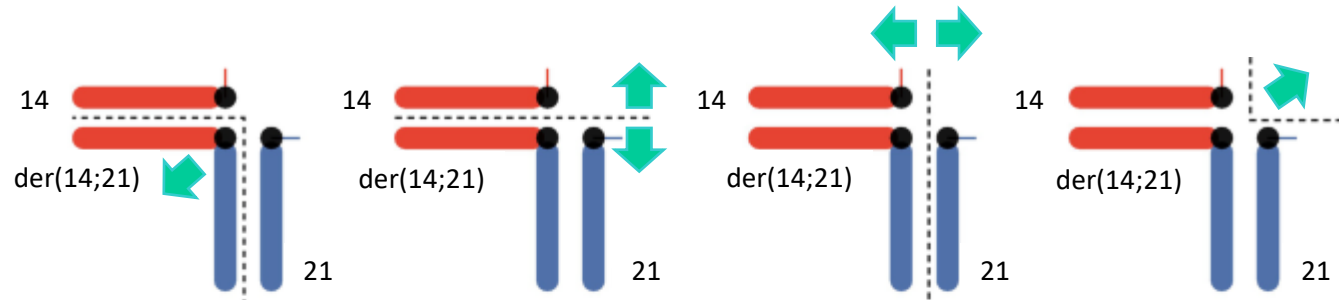


Exemple: une personne 45,XY,rob(14;21)(q10;q10) ou 45,XX,rob(14;21)(q10;q10)

Formation d'un trivalent (méiose, pachytène)

der: chromosome dérivatif
➡ sens de la ségrégation méiotique

4 possibilités de ségrégations

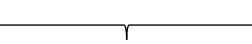
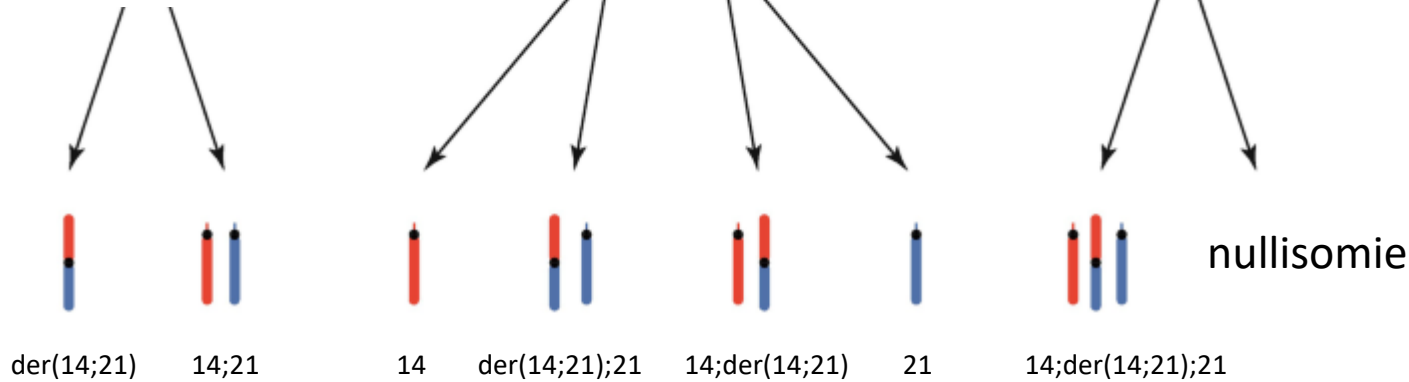


Ségrégation alterne

Ségrégation adjacente

3:0 ségrégation

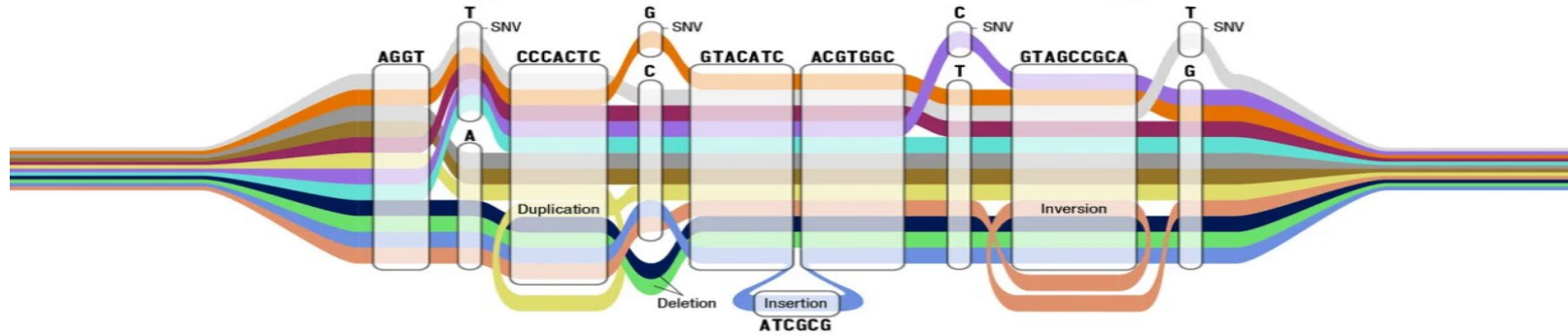
8 possibilités de gamètes



90% des cas

Conséquence possible pour la descendance ???? cf. TD de génétique

La signification clinique des variants



- Sans effet notable pour notre santé
- Expliquent nos différences

?

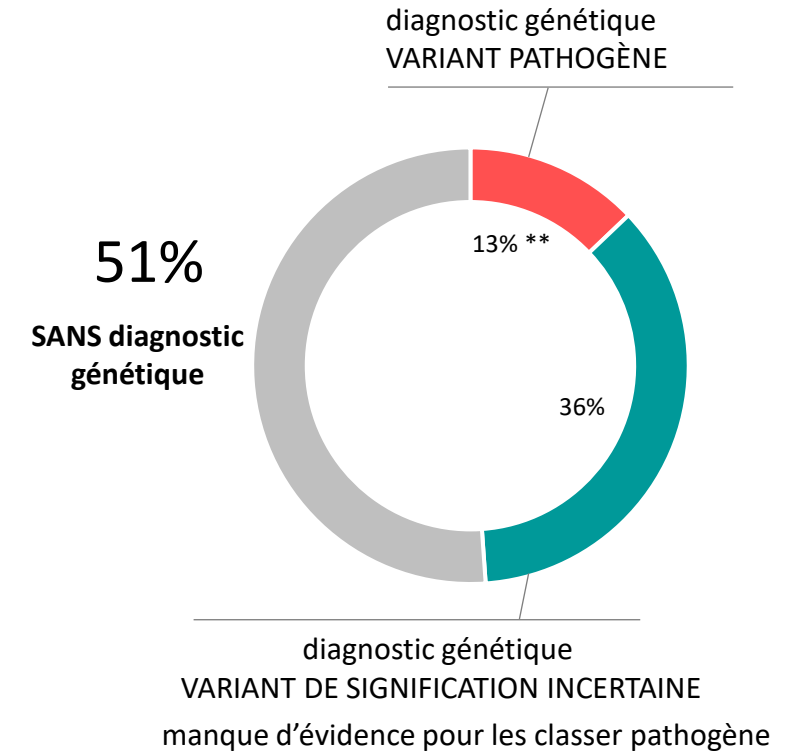
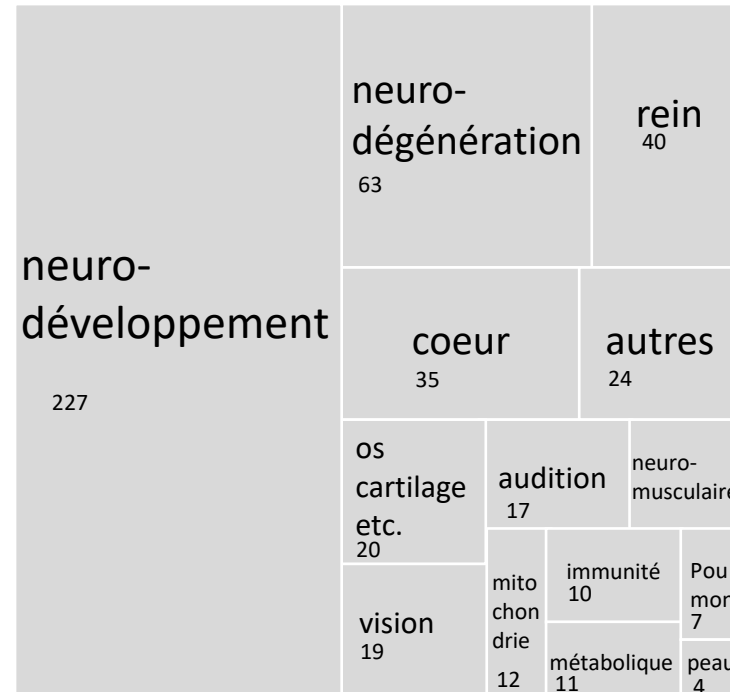
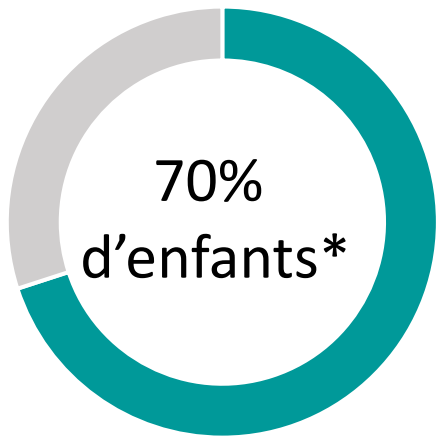


- Il a été démontré dans la littérature scientifique qu'ils causent le développement de maladies

Le diagnostic génétique des maladies



Les maladies



Objectifs de cette 3^{ème} séance

- Expliquer la fréquence allélique et son implication dans les maladies génétiques
- Citer les caractéristiques d'un variant *de novo*
- Expliquer les conséquences de la redondance du code génétique
- Différencier les variants SNV synonymes, non-synonymes
- Expliquer les variants faux-sens et ses conséquences
- Expliquer les variants non-sens et ses conséquences
- Expliquer l'origine d'un décalage de lecture et ses conséquences
- Citer les bases et sites importants pour un épissage
- Décrire un exemple d'un défaut d'épissage d'origine génétique
- Expliquer la différence d'une séquence codante et non-codante
- Décrire l'exemple d'un variant non-codant qui modifie le taux de transcription d'un gène
- Citer les éléments qui expliquent la pathogénicité d'un SV
- Expliquer l'impact des monosomies et trisomies
- Nommer les conséquences cliniques pour les porteurs/euses de translocations et leurs descendances